

Kontrast nicht gut

Wie funktioniert eigentlich Google Translate?

Utz Pöhlmann

Universität Hamburg - Proseminar: Künstliche Intelligenz

4poehlma@informatik.uni-hamburg.de

2. Juli 2015

→ Braucht Du diesen Overhead überhaupt?

Gliederung

1 Einleitung ?

- Vorschau
- Beispiel

2 Sprachübersetzungssysteme

- ~~Übersetzung von Grund auf~~

3 Schablonentechnik

4 Fazit & Ausblick

- Aussicht
- Fazit
- Quellen

nicht zurück Details
an dieser Stelle !

Seitenzahl

Warum Sprachübersetzungssysteme?

Mit dieser Folie vor der Gliederung starten?

- 7000 Sprachen weltweit

Warum Sprachübersetzungssysteme?

- 7000 Sprachen weltweit
- Reisen sind „IN“

Warum Sprachübersetzungssysteme?

- 7000 Sprachen weltweit
- Reisen sind „IN“
- Übersetzung überall und für jedermann

Ein Übersetzungsbeispiel

The screenshot shows the Google Translate web interface. At the top left is the Google logo. To its right, handwritten in red, is the text "Kannst Du das auch live machen?" followed by "(→VLAN vorher prüfen)". In the top right corner, there is a "Sign in" button and a Google+ icon. Below the logo, the word "Translate" is displayed. The interface features two language selection dropdowns: the first is set to "English" with options for Spanish, French, and "English - detected"; the second is set to "German" with options for English and Spanish. A blue "Translate" button is positioned to the right of these dropdowns. Below the language settings is a large text input area on the left and a corresponding output area on the right. At the bottom of the page, there is a banner for "Help improve Google Translate" featuring a cartoon character and a right-pointing arrow.

Ein Übersetzungsbeispiel

The screenshot shows the Google Translate web interface. At the top, the Google logo is on the left, and a 'Sign in' button is on the right. Below the logo is the 'Translate' heading. The interface shows the source language as 'English - detected' and the target language as 'German'. The input text is 'undefined', and the output translation is 'nicht definiert'. Below the input and output boxes, there are sections for 'Definitions of undefined' and 'Translations of undefined'. The definitions section includes the word 'adjective', the phrase 'not clear or defined.', the phrase 'undefined areas of jurisdiction', and synonyms: 'unspecified, unexplained, unspecific, in...'. The translations section includes the word 'adjective' and a list of German equivalents: 'undefiniert', 'nicht definiert', and 'undefinierbar', each followed by its English equivalent: 'undefined', 'undefined', and 'indefinable, undefinable' respectively. At the bottom left, there is a 'See also' section with the word 'undefine'.

Google

Sign in

Translate

g+

English Spanish French English - detected

German English Spanish

Translate

undefined

nicht definiert

Definitions of undefined

adjective

not clear or defined.

"undefined areas of jurisdiction"

synonyms: unspecified, unexplained, unspecific, in...

Translations of undefined

adjective

undefiniert undefined

nicht definiert undefined

undefinierbar indefinable, undefinable

See also

undefine

Ein Übersetzungsbeispiel

Google Translate interface showing a translation of "undefined" from English to German. The German translation is "nicht definiert". A dropdown menu is open, showing alternate translations: "nicht definiert", "undefiniert", "undefined", "undefinierten", "undefinierte", and "Improve this translation".

English Spanish German English - detected

German English Spanish Translate

Click to edit and see alternate translations

undefined

nicht definiert

nicht definiert
undefiniert
undefined
undefinierten
undefinierte
Improve this translation

Definitions of undefined

adjective
not clear or defined.
"undefined areas of jurisdiction"
synonyms: unspecified, unexplained, unspecific, in...

See also
undefined

Translations of undefined

adjective

■ undefined	undefined
■ nicht definiert	undefined
■ undefinierbar	indefinable, undefinable

Ein Übersetzungsbeispiel

The screenshot displays the Google Translate web interface. At the top left is the Google logo, and at the top right is a 'Sign in' button. Below the logo is the 'Translate' heading. On the right side of the heading are the Google+ logo and a bookmark icon. The main interface features two language selection rows. The first row has buttons for 'English', 'Spanish', 'German', and 'German - detected' with a dropdown arrow. The second row has buttons for 'German', 'English', and 'Spanish' with a dropdown arrow. A blue 'Translate' button is positioned to the right of the second row. A double-headed arrow icon is located between the two language rows. The input text box on the left contains the German sentence 'Die Weisheit jagt mich, aber ich bin schneller!' with a close button (X) in the top right corner. Below the input box are icons for a speaker (audio) and a pencil (edit). The output text box on the right contains the English translation 'The wisdom of chasing me, but I am faster!'. Below the output box are icons for a star (favorites), a list (history), a speaker (audio), and a 'Wrong?' button with a pencil icon.

Google

Sign in

Translate

g+

English Spanish German German - detected

↔

German English Spanish

Translate

Die Weisheit jagt mich, aber ich bin schneller! ✕

The wisdom of chasing me, but I am faster!

Wrong?

Ein Übersetzungsbeispiel

The screenshot displays the Google Translate web interface. At the top, the Google logo is on the left, and a 'Sign in' button is on the right. Below the logo, the word 'Translate' is prominently displayed. To the right of 'Translate' are the Google+ logo and a bookmark icon. The main area features two language selection dropdowns: the first is set to 'German - detected' and the second to 'English'. A blue 'Translate' button is positioned to the right of these dropdowns. A tooltip with the text 'Click to edit and see alternate translations' points to the 'Translate' button. The input text on the left is 'Die Weisheit jagt mich, aber ich bin schneller!', with 'Die Weisheit' highlighted in blue. The output text on the right is 'The wisdom of chasing me, but I am faster!', with 'The wisdom of' highlighted in blue. Below the input text are icons for a speaker and an edit pencil. Below the output text are icons for a star, a list, a speaker, and a 'Wrong?' link with an edit pencil icon.

Google

Sign in

Translate

g+

English Spanish German German - detected

German English Spanish Translate

Click to edit and see alternate translations

Die Weisheit jagt mich, aber ich bin schneller!

The wisdom of chasing me, but I am faster!

Wrong?

Ein Übersetzungsbeispiel

The screenshot displays the Google Translate web interface. At the top, the Google logo is on the left, and a 'Sign in' button is on the right. Below the logo, the word 'Translate' is prominently displayed. The interface features two input fields for text. The left field contains the German sentence 'Die Weisheit jagt mich, aber ich bin schneller!'. The right field contains the English translation 'The wisdom of chasing me, but I'. A dropdown menu is open under the word 'The' in the English translation, showing suggestions: 'The wisdom of', 'The wisdom', 'Wisdom', 'Die Weisheit', and 'Improve this translation'. The word 'Weisheit' is also visible in a separate box to the right of the dropdown. At the bottom of the interface, there are icons for audio playback and a 'Wrong?' link for feedback.

Google

Sign in

Translate

g+

English Spanish German German - detected

German English Spanish

Translate

Die Weisheit jagt mich, aber ich bin schneller!

The wisdom of chasing me, but I

The wisdom of
The wisdom
Wisdom
Die Weisheit
Improve this translation

Weisheit

Wrong?

Ein Übersetzungsbeispiel

The screenshot displays the Google Translate web interface. At the top left is the Google logo, and at the top right is a 'Sign in' button. Below the logo is the 'Translate' heading. The main input area shows a German sentence: 'Die Weisheit jagt mich, aber ich bin schneller!'. The source language is set to 'German - detected'. The target language is set to 'English'. The translated text is 'Wisdom chasing me, but I am faster!'. Below the input and output boxes are various icons for voice input, editing, and feedback. A yellow banner at the bottom of the interface states: 'Your contribution will be used to improve translation quality and may be shown to users anonymously'. At the very bottom, there is a search bar with the word 'chasing' and a 'See also' link.

Google

Sign in

Translate

g+

English Spanish German German - detected

German English Spanish

Translate

Die Weisheit jagt mich, aber ich bin schneller!

Wisdom chasing me, but I am faster!

Wrong? Undo edits

Your contribution will be used to improve translation quality and may be shown to users anonymously

Contribute Close

chasing

See also

Ein Übersetzungsbeispiel

The screenshot shows the Google Translate web interface. At the top left is the Google logo, and at the top right is a 'Sign in' button. Below the logo is the 'Translate' heading. The interface has two main text input areas. The left area contains the German text 'Die Weisheit jagt mich, aber ich bin schneller!'. The right area contains the English translation 'Wisdom chasing me, but I am faster!'. A dropdown menu is open over the word 'chasing' in the English text, showing suggestions: 'chasing', 'chases', 'hunts', 'hunting', 'hunt', and 'Improve this translation'. To the right of the English text, there is a button for 'Verfolgungsjagden' and a 'Wrong? Undo edits' link. Below the English text, there is a yellow banner that reads: 'Your contribution will be used to improve translation quality and may be shown to users anonymously'. Below the banner are two buttons: 'Contribute' and 'Close'. At the bottom left, there is a small audio icon and the text 'See also'. At the bottom right, there are navigation arrows and a close button.

Google

Translate

English Spanish German German - detected

German English Spanish

Translate

Die Weisheit jagt mich, aber ich bin schneller!

Wisdom chasing me, but I am faster!

chasing
chases
hunts
hunting
hunt
Improve this translation

Verfolgungsjagden

Wrong? Undo edits

Your contribution will be used to improve translation quality and may be shown to users anonymously

Contribute Close

chasing

See also

Ein Übersetzungsbeispiel

The screenshot displays the Google Translate web interface. At the top left is the Google logo, and at the top right is a 'Sign in' button. Below the logo is the 'Translate' heading. On the right side of the heading are the Google+ logo and a bookmark icon. The main interface features two language selection dropdowns: the first is set to 'German - detected' and the second to 'English'. A 'Translate' button is positioned to the right of these dropdowns. The input text box on the left contains the German sentence 'Die Weisheit jagt mich, aber ich bin schneller!'. The output text box on the right shows the English translation 'Wisdom chases me, but I am faster!'. Below the German text are icons for audio playback and editing. Below the English text are icons for a star, a list, audio playback, and links for 'Wrong?' and 'Undo edits'. A yellow banner at the bottom of the output area states: 'Your contribution will be used to improve translation quality and may be shown to users anonymously'. Below this banner are 'Contribute' and 'Close' buttons. At the very bottom, the word 'chases' is shown with an audio icon, and navigation arrows are visible on the right.

Google

Sign in

Translate

g+

English Spanish German German - detected

↔

German English Spanish

Translate

Die Weisheit jagt mich, aber ich bin schneller!

Wisdom chases me, but I am faster!

Wrong? Undo edits

Your contribution will be used to improve translation quality and may be shown to users anonymously

Contribute Close

chases

See also

Ein Übersetzungsbeispiel

The screenshot displays the Google Translate web interface. At the top left is the Google logo, and at the top right is a 'Sign in' button. Below the logo is the 'Translate' heading. The interface shows a source text box on the left containing the German sentence 'Die Weisheit jagt mich, aber ich bin schneller!' with a close button (X) and audio/play icons. The target text box on the right shows the English translation 'Wisdom chases me, but I am faster!' with a star icon, a list icon, an audio icon, and links for 'Wrong?' and 'Undo edits'. Below the translation boxes is a yellow banner that reads 'Thank you for your submission.' At the bottom, there is a section for 'chases' with an audio icon, and a 'See also' section listing 'chase'. Navigation arrows and a close button are visible on the right side of the bottom section.

Google

Sign in

Translate

g+

English Spanish German German - detected

↔

German English Spanish

Translate

Die Weisheit jagt mich, aber ich bin schneller! ✕

Wisdom chases me, but I am faster!

☆ ☰ 🔊

✎ Wrong? ↶ Undo edits

Thank you for your submission.

chases 🔊

See also

chase

< > ✕

Arten von Übersetzungssystemen

- Wort für Wort

Arten von Übersetzungssystemen

- Wort für Wort
- Schablonentechnik

Arten von Übersetzungssystemen

- Wort für Wort
 - Übersetzen Marke Lexikon
- Schablonentechnik

Arten von Übersetzungssystemen

- Wort für Wort
 - Übersetzen Marke Lexikon
 - Sinnabschnitte und Wahrscheinlichkeiten
- Schablonentechnik

Oh, aber vermeide:
– leere Folien
– zu viele Klicks

Übersetzen Marke Lexikon

- Gehe zum ersten Wort

Übersetzen Marke Lexikon

- Gehe zum ersten Wort
- Schau ins Lexikon

Übersetzen Marke Lexikon

- Gehe zum ersten Wort
- Schau ins Lexikon
- Nimm die erste Wortbedeutung

Übersetzen Marke Lexikon

- Gehe zum ersten Wort
- Schau ins Lexikon
- Nimm die erste Wortbedeutung
- Nächstes Wort

Übersetzen Marke Lexikon

- Gehe zum ersten Wort
- Schau ins Lexikon
- Nimm die erste Wortbedeutung
- Nächstes Wort
- Fahre fort mit Schritt 2

Nutze
den
Space,
... Luke

Sinnabschnitte und Wahrscheinlichkeiten

- Parallelen

Sinnabschnitte und Wahrscheinlichkeiten

- Parallelen
- Worttypen

Sinnabschnitte und Wahrscheinlichkeiten

- Parallelen
- Worttypen
- Voraussagen

Sinnabschnitte und Wahrscheinlichkeiten

- Parallelen
- Worttypen
- Voraussagen
- Filterung

Sinnabschnitte und Wahrscheinlichkeiten

- Parallelen
- Worttypen
- Voraussagen
- Filterung
- Übersetzen

↓
Warum nicht parallel am Beispiel
zeigen?

Schablonentechnik

- Textabschnitte in anderen Texten wiederfinden

Schablonentechnik

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen

Schablonentechnik

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen
- Als Schablone behalten

↓ ebenfalls:
gleich an Beispiel
aufbauen !

Schablonentechnik

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen
- Als Schablone behalten
- Schablonen, die am wenigsten in Konflikt stehen nehmen

Schablonentechnik

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen
- Als Schablone behalten
- Schablonen, die am wenigsten in Konflikt stehen nehmen
- Diese Schablonen als aktuelle Übersetzung ausgeben

Aussicht

- Diverse Versuche, bessere Übersetzungssysteme herzustellen

Aussicht

| Vermeide Meta-talk
⇒ Konkrete Probleme / Fakten
⇒ Sachliche Gründe

- Diverse Versuche, ~~bessere~~ Übersetzungssysteme herzustellen
ble
- Systeme aber ~~bisher~~ immer schlechter als Google Translate
Genauer? Fakten & Zahlen?
Empirischer Versuch?

Fazit

- Internetübersetzungssysteme recht gut

Fazit

- Internetübersetzungssysteme recht gut
- Wird ständig verbessert

?

Siehe letzte Folie!
→ konkrete Fakten,
vermeide hohle Phrasen

Fazit

- Internetübersetzungssysteme recht gut
- Wird ständig verbessert
- Gänzlich neue Lösung jedoch nicht in Sicht

Quellenverzeichnis I

Fokussiere auf die wichtigsten
2-5 Quellen \Rightarrow nur eine Seite



Kamal Nicholas (2015)

Google Translate: Über 80 Sprachen verstehen

<http://www.giga.de/webapps/google-translate/> Absatz 3, Mai 2015.



Adak, C. (2014)

A bilingual machine translation system: English and Bengali

Automation, Control, Energy and Systems (ACES), 2014 First International Conference on pages 1 – 4, Feb 2014.



Brkic, M. and Mikulic, B.B. and Matetic, M. (2012)

Can we beat google translate?

Information Technology Interfaces (ITI), Proceedings of the ITI 2012 34th International Conference on pages 381 – 386, June 2012.

Quellenverzeichnis II



Klyuev, V. and Haralambous, Y. (2012)

Query translation for CLIR: EWC vs. Google Translate

Information Science and Technology (ICIST), 2012 International Conference on
pages 707 – 711, March 2012.



Miltsakaki, Eleni and Robaldo, Livio and Lee, Alan and Joshi, Aravind (2008)

Sense annotation in the penn discourse treebank

Computational Linguistics and Intelligent Text Processing pages 275 – 286,
Springer, 2008.



Wan, Xiaojun (2008)

Using bilingual knowledge and ensemble techniques for unsupervised Chinese
sentiment analysis

*Proceedings of the Conference on Empirical Methods in Natural Language
Processing* pages 553 – 561. Association for Computational Linguistics, 2008.

Quellenverzeichnis III



Demberg, Vera (2007)

A language-independent unsupervised model for morphological segmentation
ANNUAL MEETING-ASSOCIATION FOR COMPUTATIONAL LINGUISTICS
volume 45, page 920. Citeseer, 2007.



Toutanova, Kristina and Ilhan, H Tolga and Manning, Christopher D (2002)

Extensions to HMM-based statistical word alignment models
Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10 pages 87 – 94. Association for Computational Linguistics, 2002.



Kaji, Hiroyuki and Kida, Yuuko and Morimoto, Yasutsugu (1992)

Learning translation templates from bilingual text
Proceedings of the 14th conference on Computational linguistics-Volume 2 pages 672 – 678. Assosiacion of Computational Linguistics, 1992.

Ende

Ende

Noch Fragen?

Krasser Stuff der mir eventuell noch einfällt.