



LT5: Software-Unterstützung für Korrektheit und Leichte Sprache

Tekom Jahrestagung 2013

7. November 2013

Daniel Naber

Annika Nietzio



Überblick

1. Was ist LanguageTool?
2. Wie findet LanguageTool Fehler?
3. Wie kann man LanguageTool erweitern?
4. Was ist Leichte Sprache?
5. Wie kann man mit LanguageTool Leichte Sprache prüfen?

Über die Vortragenden

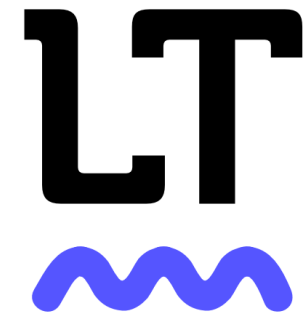
- **Annika Nietzio**

- Forschungsinstitut Technologie und Behinderung, Wetter (Ruhr)
- Forschung und Beratung zu barrierefreier Kommunikation
- Büro für Leichte Sprache Volmarstein



- **Daniel Naber**

- Computerlinguist, Potsdam
- Entwickler von LanguageTool





LanguageTool

Finde die Fehler

“ Berthold Brecht ist tod?
Das kann nicht seien!

Finde die Fehler

Berthold Brecht ist tod? Dann kann nicht seien!

Meinen Sie den deutschen Dramatiker 'Bertolt' Brecht? (Das ist die von ihm selbst bevorzugte Schreibweise des Vornamens.)

Bertolt

Hier ignorieren

Fehler dieses Typs ignorieren

Finde die Fehler

Berthold Brecht ist tod? Dann kann nicht seien!

Meinten Sie den Infinitiv 'sein'?

sein

Hier ignorieren

Fehler dieses Typs ignorieren

LanguageTool

- Open-Source-Software zur Textprüfung
- seit 2003
- unterstützt u.a. Deutsch, Englisch, Französisch, Spanisch
- leicht erweiterbar

Open Source bedeutet hier

- kostenlos
- herunterladen ohne Registrierung
- Regeln (XML) und Sourcecode (Java) frei verfügbar und änderbar
- Lizenz: LGPL - auch kommerziell nutzbar und in kommerzielle Software integrierbar

LanguageTool-Ansatz

- LanguageTool fängt da an, wo die klassische Rechtschreibprüfung aufhört
- klassische Rechtschreibprüfung = prüft einzelne Wörter, ohne Kontext
- Rechtschreibung über externe Komponente (hunspell)

LanguageTool-Prinzip

- Text wird analysiert
 - Erkennung von Satzgrenzen, Wörtern und Wortarten
- Analysierter Text wird auf Fehlermuster durchsucht
 - LanguageTool-Fehlermuster kennen keine Grammatik, sie kennen Fehler.
- komplizierte Regeln in Java programmiert

Fehlermuster-Beispiel

```
<rule>
  <pattern>
    <token>Berthold</token>
    <token>Brecht</token>
  </pattern>
  <message>Meinen Sie den deutschen
    Dramatiker
    <suggestion>Bertolt Brecht</suggestion>?
  </message>
</rule>
```

Was können Fehlermuster?

- Reguläre Ausdrücke:
`<token regexp="yes">Berth?old</token>`
- Wortarten, z.B. alle Verben:
`<token postag_regexp="yes"
 postag="VER:.*" />`
- Negation:
`<token negate="yes">Haus</token>`

Finde den Fehler

“ Mittelohrenzündung

Finde den Fehler

Mittelohrenzündung =
Mittel + Ohren + Zündung ?

Es wurde eine **Mittelohrenzündung** diagnostiziert.

Meinten Sie 'Mittelohrenzündung'?

Mittelohrenzündung

Hier ignorieren

Fehler dieses Typs ignorieren

Finde den Fehler

“ Weiterbildungen gemäß des BKrFQG

Finde den Fehler

Weiterbildungen **gemäß des** BKrFQG

Meinten Sie 'gemäß dem'? Die Präposition 'gemäß' erfordert den Dativ.

gemäß dem

Mehr Informationen...

Hier ignorieren

Fehler dieses Typs ignorieren

Grenzen der Fehlererkennung

- Fehlalarme: „Brecht“

Berthold Brecht ist tod? Das kann nicht seien!

- Nicht über Muster abzudecken: Fehler über Satzgrenzen
- Schwer über Muster abzudecken: fehlende Wörter
- deshalb: komplizierte Regeln sind in Java programmiert

Nutzung von LanguageTool

- übers Web <http://languagetool.org/de>
- als Desktop-Applikation (Java)
- als Firefox-Add-on
- als LibreOffice/OpenOffice-Add-on
- für Entwickler:
 - über Java-API
 - über HTTP-Schnittstelle, die XML zurückliefert

LanguageTool Varianten

- LanguageTool ist erweiterbar ohne Programmierkenntnisse...
 - 29 unterstützte Sprachen
 - Simplified Technical English (ASD-STE100), extern
 - Leichte Sprache



Leichte Sprache

Im UN-Vertrag steht:



Die Informationen müssen zugänglich sein.

Die Informationen müssen so sein,
dass Menschen mit Behinderung
die Informationen verstehen können.

Zum Beispiel muss es Informationen
in Blinden-Schrift oder in Leichter Sprache geben.

Besonders Ämter und Behörden müssen
zugängliche Informationen machen.

Auch das Radio, Fernsehen, die Zeitung
und das Internet sollen zugänglich sein.

Im UN-Vertrag steht:



Die Informationen müssen zugänglich sein.

Die Informationen müssen so sein,
dass Menschen mit Behinderung
die Informationen verstehen können.

Zum Beispiel muss es Informationen
in Blinden-Schrift oder in Leichter
Schrift geben.

Besonders Ämter und Behörden
müssen diese Informationen
zugängliche Informationen machen.

Auch das Radio, Fernsehen, die Zeitung
und das Internet sollen zugänglich sein.

**Kurze Sätze.
Meist nur Hauptsätze.
Höchstens ein Nebensatz.**



Im UN-Vertrag steht:

Die Informationen müssen zugänglich sein.

Die Informationen müssen so sein,
dass Menschen mit Behinderungen
die Informationen verstehen können.

Zum Beispiel muss es Informationen
in **Blinden-Schrift** oder in Leichter
Schrift geben.

Besonders **Ämter** und Behörden müssen
zugängliche Informationen machen.

Auch das Radio, Fernsehen, die Zeitung
und das Internet sollen zugänglich sein.

Kurze Wörter.
Zusammengesetzte Wörter
mit Bindestrich schreiben.

Im UN-Vertrag steht:



Die **Informationen** müssen zugänglich sein.

Die **Informationen** müssen so sein,
dass Menschen mit Behinderung
die **Informationen** verstehen können.

Zum Beispiel muss es **Informationen**
in Blinden-Schrift oder in Leichter Sprache geben.

Besonders Ämter und Behörden
zugängliche **Informationen** machen.

Auch das Radio, Fernsehen, die
und das Internet sollen zugänglich sein.

Das gleiche Wort für die
gleiche Sache verwenden.
Wenig Verwendung
von Pronomen.

Im UN-Vertrag steht:



Die Informationen müssen zugänglich sein.

Die Informationen müssen so sein,
dass Menschen mit Behinderung
die Informationen verstehen können.

Zum Beispiel muss es Informationen
in Blinden-Schrift oder in Leichter Sprache geben.

Besonders Ämter und Behörden müssen
zugängliche Informationen machen.

Beispiele

**Auch das Radio, Fernsehen, die Zeitung
und das Internet** sollen zugänglich sein.

Im UN-Vertrag steht:



Die Informationen müssen zugänglich sein.

Die Informationen müssen so sein,
dass Menschen mit Behinderung
die Informationen verstehen können.

Zum Beispiel muss es Informationen geben, die
in Blinden-Schrift oder in Leichter Sprache
angeboten werden. **Texte von der Zielgruppe
prüfen lassen.**

Besonders Ämter und Behörden müssen
zugängliche Informationen machen.

Auch das Radio, Fernsehen, die Zeitung
und das Internet sollen zugänglich sein.

Regeln für Leichte Sprache

- Netzwerk Leichte Sprache:
Regeln für Leichte Sprache
- Inclusion Europe: **Informationen für alle:**
Europäische Regeln, wie man Informationen leicht lesbar und leicht verständlich macht.
 - in 12 Sprachen
- **Barrierefreie Informationstechnik-Verordnung (BITV) 2.0**
 - Leichte Sprache in Webangeboten des Bundes

Woher kommt Leichte Sprache?

- Veränderung der Sichtweise:
 - Auch Menschen mit Lernschwierigkeiten und geistigen Behinderungen können und sollen verstehen, worum es geht.
- Selbstvertretungsbewegung „People first“
- Politische Forderung „Halt! Leichte Sprache!“
 - UN-Behindertenrechtskonvention
- Menschen mit Lernschwierigkeiten Behinderungen arbeiten mit:
 - als Prüfer
 - bei Schulungen



Zielgruppen

- Menschen mit Lernschwierigkeiten, Menschen mit geistigen Behinderungen
- Menschen mit geringer Schriftsprachkompetenz, funktionale Analphabeten
- Menschen mit geringen Deutschkenntnissen.

Studie *Easy Web*, Österreich 2009

Informationen in Leichter Sprache:

- für 12% der Bevölkerung sinnvoll.
- für 4% der Bevölkerung notwendig.

Wie schreibt man in Leichter Sprache?

- Vor dem Schreiben: Analyse des Ausgangstexts.
 - Welche Informationen sind für die Leser wichtig?
 - Wie kann ich einen Bezug zur Situation der Leser herstellen?
 - Welche Wörter müssen erklärt werden, weil die Leser sie nicht kennen?
- Wenn der Textentwurf fertig ist:
 - Kunde prüft den Inhalt
 - Zielgruppe prüft die Verständlichkeit
 - Technische Unterstützung zur Prüfung der Regeln für Leichte Sprache

Beispiel: Ausgangstext

Eine wichtige Voraussetzung für gleichberechtigte Teilhabe und unabhängige Lebensführung von Menschen mit Behinderungen sind barrierefreie Informationen. Denn diese Informationen brauchen die Menschen für ein selbstbestimmtes Leben. Ziel des Aktionsplans ist, dass alle Menschen in Wetter gut verständliche Informationen bekommen.

Der Bedarf an Leichter Sprache nimmt zu, weil Menschen mit Behinderungen zunehmend selbständiger leben, z.B. in einer eigenen Wohnung. Darum sollten auch Stromrechnungen, Mietverträge und Ähnliches Erklärungen in Leichter Sprache enthalten.

Beispiel: Text in Leichter Sprache (1. Fassung)

Alle Menschen in Wetter sollen gute und verständliche Informationen bekommen. Dann können die Menschen verstehen, worum es geht. Und mitreden.

Darum müssen Informationen in Leichter Sprache zur Verfügung gestellt werden.

Leben Menschen mit Behinderung in einer eigenen Wohnung, müssen sie viele Dinge lesen.

Dann ist Leichte Sprache besonders wichtig, z.B. ein Miet-Vertrag in Leichter Sprache, Briefe der Behörden in Leichter Sprache.

Beispiel: Text in Leichter Sprache (mit **LanguageTool-Prüfergebnissen**)

Alle Menschen in Wetter sollen gute und verständliche Informationen bekommen. Dann können die Menschen verstehen, worum es geht. Und mitreden.

Darum müssen Informationen in Leichter Sprache zur Verfügung gestellt werden.

Leben Menschen mit Behinderung in einer eigenen Wohnung, müssen sie viele Dinge lesen.

Dann ist Leichte Sprache besonders wichtig, z.B. ein Miet-Vertrag in Leichter Sprache, Briefe der Behörden in Leichter Sprache.

Beispiel: Text in Leichter Sprache (Überarbeitete Fassung)

Alle Menschen in Wetter sollen gute und verständliche Informationen bekommen.

Dann können die Menschen verstehen, worum es geht. Und mitreden.

Darum muss es Informationen in Leichter Sprache geben.

Wenn Menschen mit Behinderung in einer eigenen Wohnung leben, müssen sie viele Dinge lesen.

Dann ist Leichte Sprache besonders wichtig.

Zum Beispiel:

- ein Miet-Vertrag in Leichter Sprache
- Briefe vom Amt in Leichter Sprache

Fehlermuster: Genitiv

```
<rule>
  <pattern>
    <token postag_regexp="yes" postag="SUB:.*" />
    <token postag_regexp="yes"
      postag="ART:(DEF|IND):GEN:.*" skip="-1" />
    <token postag_regexp="yes"
      postag="SUB:GEN:.*" />
  </pattern>
  <message>Genitiv gefunden: <match no="2" />
    Vermeiden Sie den Genitiv.</message>
</rule>
```

Fazit

- Was ist LanguageTool?
 - Eine Open-Source-Software zur Textprüfung
- Wie findet LanguageTool Fehler?
 - Indem es im Text nach Fehlermustern sucht
- Wie kann man LanguageTool erweitern?
 - Indem man eigene Fehlermuster in XML schreibt
- Was ist Leichte Sprache?
 - Leichte Sprache ist für Menschen mit Lernschwierigkeiten.
 - Für Leichte Sprache gibt es Regeln für Sprache, Inhalt und Präsentation.
- Wie kann man mit LanguageTool Leichte Sprache prüfen?
 - Indem man Fehlermuster für die Regeln der Leichten Sprache schreibt



Fragen!

und Antworten?

Links (1)

- LanguageTool online ausprobieren:
<http://languagetool.org/de>
- LanguageTool für Leichte Sprache online ausprobieren:
<http://languagetool.org/de/leichte-sprache/>
- Lernen, selber Regeln zu schreiben:
<http://languagetool.org/development/#intro>

Links (2)

- Die Regeln für Leichte Sprache:
<http://www.leichtesprache.org>
- Informationen für alle : Europäische Regeln, wie man Informationen leicht lesbar und leicht verständlich macht.
<http://inclusion-europe.org>
- Barrierefreie Informationstechnik-Verordnung (BITV) 2.0: Anlage 2 mit Verweisen auf den Text der Begründung zur BITV 2.0.
<http://www.di-ji.de/r/bitvanlage2>