



Erkennung von Stilrichtlinien- verletzungen in Gesetzestexten

Linguistisches Forschungskolloquium HS 2011

Stefan Höfler

Motivation

- Gesetzestexte müssen **formalen und sprachlichen Anforderungen** entsprechen.
- Diese Anforderungen sind im Bund und in vielen Kantonen in ausführlichen **Richtlinien** festgehalten.
- Im Bund überprüft die **Zentralen Sprachdienste der Bundeskanzlei** Gesetzesentwürfe auf Einhaltung der sprachlichen Richtlinien.

Problem

- Die Überprüfung eines Gesetzesentwurfs «von Hand» ist **aufwändig**.

Projekt

- Entwicklung von Methoden, um bestimmte **Richtlinienverletzungen** **maschinell erkennen** zu können

Ausgangspunkt: technische Redaktion

Firmen verfügen oft über **interne Stilregeln** für ihre technische Dokumentation (Benutzerhandbücher etc.), z.B.

- kein Passiv verwenden
- Titel nicht mit einem Artikel beginnen
- kein Futur verwenden
- Massangaben abkürzen
- ...

Maschinelle Stilprüfung

Autorenwerkzeuge

- **markieren** Textpassagen, die eine Stilregel verletzen
- **informieren** den Benutzer darüber, welche Stilregel in der Passage (potentiell) verletzt wurde

Beispiele

- **CLAT** von Congree Language Technologies (www.congree.com)
- **acrolinx IQ** von acrolinx (www.acrolinx.de)

Vorgehen

- maschinelle **Vorverarbeitung** des Textes
- maschinelle **Suche** nach Regelverletzungen im vorverarbeiteten Text

Vorverarbeitung

Textentwurf wird maschinell analysiert und mit der so gewonnenen **strukturellen und linguistischen Information** angereichert:

- **Tokenisierung**
maschinelle Wortgrenzenerkennung
- **Textsegmentierung**
maschinelles Erkennen von Kapitelgrenzen, Abschnittsgrenzen, Überschriften, Fussnoten, Satzgrenzen, ...
- **Part-of-speech Tagging**
maschinelle Wortartenerkennung
- **Morphologische Analyse**
maschinelles Erkennen von Kasus, Genus, Numerus, ...
- **Syntaktische Analyse (Parsing)**
maschinelles Erkennen der Satzstruktur

Suche nach Richtlinienverletzungen

Spezielle Suchregeln beschreiben die exakten Bedingungen, unter denen eine bestimmte Stilrichtlinie *verletzt* ist («**Fehlermodellierung**»).

Avoid_future

Die Stilregel «Futur vermeiden» ist (unter anderem) verletzt, wenn...

```
/* Example: ".. It will be necessary .." */
```

```
TRIGGER (80)==@will1 [-@comma]* @verbInf2
```

```
->($will, $verbInf)
```

```
->{mark : $will, $verbInf;}
```

... eine Wortform des Lemmas *will* von einem Infinitiv gefolgt wird und dazwischen null, ein oder mehrere Token sind, aber kein Komma.

In diesem Fall markiere *will* und den Infinitiv – und gib Fehlermeldung 80 aus.

Quelle

Lehmann, S. (2009) Kontrollierte Sprachen und Sprach-technologie in der Industrie: das Autorenwerkzeug acrolinx. UZH. https://cast.switch.ch/vod/clips/2pjsz3usia/link_box

Fehlermodellierung

Die Methode der Fehlermodellierung anzuwenden, bedeutet also,

- anstatt die «Wohlgeformtheit» eines ganzen Textes zu überprüfen,
- zu **antizipieren (bzw. «modellieren»)**, wie spezifische Richtlinienverletzungen aussehen, und gezielt nach solchen zu suchen.

Anmerkung «Fehlermodellierung»

Richtlinien sind meist keine absoluten Stilregeln, sondern oft «nur» **Empfehlungen** oder **Faustregeln**.

Gesucht wird also nicht nach eigentlichen «Fehlern», sondern nach **potentiellen** Richtlinienverletzungen.

Geht das auch für die Gesetzesredaktion?

² Über Rechtsfragen von grundsätzlicher Bedeutung oder auf Antrag eines Richter oder einer Richterin entscheiden sie in Fünferbesetzung. Ausgenommen sind Beschwerden gegen Entscheide der kantonalen Aufsichtsbehörden in Schuldbetreibungs- und Konkursachen.

³ In Fünferbesetzung entscheiden sie ferner über Beschwerden gegen referendumspflichtige kantonale Erlasse und gegen kantonale Entscheide über die Zulässigkeit einer Initiative oder das Erfordernis eines Referendums. Ausgenommen sind Beschwerden, die eine Angelegenheit einer Gemeinde oder einer anderen Körperschaft des kantonalen Rechts betreffen.

Art. 21 Abstimmung

¹ Das Gesamtgericht, die Präsidentenkonferenz, die Verwaltungskommission und die Abteilungen treffen die Entscheide, Beschlüsse und Wahlen, wenn das Gesetz nichts anderes bestimmt, mit der absoluten Mehrheit der Stimmen.

Geht das auch für die Gesetzesredaktion?

² Über Rechtsfragen von grundsätzlicher Bedeutung oder auf Antrag eines Richter oder einer Richterin entscheiden sie in Fünferbesetzung. Ausgenommen sind Beschwerden gegen Entscheide der kantonalen Aufsichtsbehörden in Schuldbetreibungs- und Konkursachen.

³ In Fünferbesetzung entscheiden sie ferner über Beschwerden gegen referendumpflichtige kantonale Erlasse und gegen kantonale Entscheide über die Zulässigkeit einer Initiative oder das Erfordernis eines Referendums. Ausgenommen sind Beschwerden, die eine Angelegenheit einer Gemeinde oder einer anderen Körperschaft des kantonalen Rechts betreffen.

Art. 21 Abstimmung

¹ Das Gesamtgericht, die Präsidentenkonferenz, die Verwaltungskommission und die Abteilungen treffen die Entscheide, Beschlüsse und Wahlen, wenn das Gesetz nichts anderes bestimmt, mit der absoluten Mehrheit der Stimmen.

Kommentar

Komplexe Koordinationen sind oft besser lesbar, wenn sie in eine Aufzählung (Bst. a, b, ...) aufgegliedert werden.

Weitere Informationen: [GTR Rz.187 ff.](#); ...

Beispiel: ...

Geht das auch für die Gesetzesredaktion?

Herausforderung für die Vorverarbeitung

Gesetzessprache

- komplex
 - idiosynkratisch
- ➔ Methoden für die Vorverarbeitung müssen auf die Eigenheiten der Gesetzessprache angepasst werden.

Herausforderung für die Fehlermodellierung

sprachliche Richtlinien für Gesetzestexte

- domänenspezifisch
 - abstrakt
- ➔ Es braucht eine spezielle Fehlermodellierung für Richtlinienverletzungen in Gesetzestexten.

Typen sprachlicher Richtlinien für Gesetzestexte

Zu bestimmten Ausdrücken

- Das Wort *beziehungsweise* darf nicht abgekürzt werden.
- Vorwärtsverweise sind nicht erwünscht.

Zu syntaktischen Konstruktionen

- Mehrfach verschachtelte Sätze sollen vermieden werden.
- Komplexe Koordinationen sollen in explizite Aufzählungen aufgeteilt werden.

Zur Diskursstruktur

- Argumentative Aussagen (Erklärungen, Rechtfertigungen) gehören nicht in einen Gesetzestext.
- Geltungsbereichsbestimmungen und Legaldefinitionen sollen keine materiellen Aussagen machen.

Die Eugen-Huber-Regel

Stilrichtlinie

- Pro Artikel höchstens drei Absätze.
- Pro Absatz ein Satz.



Eugen Huber (1849–1923),
Autor des Schweizerischen
Zivilgesetzbuchs (ZGB)

Beispiel: Bundesgerichtsgesetz (BGG)

2. Kapitel: Allgemeine Verfahrensbestimmungen

1. Abschnitt: Zuständigkeit

Art. 29 Prüfung

¹ Das Bundesgericht prüft seine Zuständigkeit von Amtes wegen.

² Bestehen Zweifel, ob das Bundesgericht oder eine andere Behörde zuständig ist, so führt das Gericht mit dieser Behörde einen Meinungs austausch.

Art. 30 Unzuständigkeit

¹ Erachtet sich das Bundesgericht als nicht zuständig, so tritt es auf die Sache nicht ein.

² Hat sich in einem Meinungs austausch die Zuständigkeit einer anderen Behörde ergeben oder erscheint die Zuständigkeit einer anderen Bundesbehörde als wahrscheinlich, so überweist das Bundesgericht die Sache der betreffenden Behörde.

Art. 31 Vorfragen

Ist das Bundesgericht in der Hauptsache zuständig, so befindet es auch über die Vorfragen.

Beispiel: Bundesgerichtsgesetz (BGG)

Art. 42 Rechtsschriften

¹ Rechtsschriften sind in einer Amtssprache abzufassen und haben die Begehren, deren Begründung mit Angabe der Beweismittel und die Unterschrift zu enthalten.

² In der Begründung ist in gedrängter Form darzulegen, inwiefern der angefochtene Akt Recht verletzt. Ist eine Beschwerde nur unter der Voraussetzung zulässig, dass sich eine Rechtsfrage von grundsätzlicher Bedeutung stellt oder ein besonders bedeutender Fall nach Artikel 84 vorliegt, so ist auszuführen, warum die jeweilige Voraussetzung erfüllt ist.

³ Die Urkunden, auf die sich die Partei als Beweismittel beruft, sind beizulegen, soweit die Partei sie in Händen hat; richtet sich die Rechtsschrift gegen einen Entscheid, so ist auch dieser beizulegen.

⁴ Bei elektronischer Zustellung muss das Dokument, das die Rechtsschrift und die Beilagen enthält, von der Partei oder ihrem Vertreter beziehungsweise ihrer Vertreterin mit einer anerkannten elektronischen Signatur versehen werden. Das Bundesgericht bestimmt in einem Reglement, in welchem Format die elektronische Zustellung erfolgen kann.

⁵ Fehlen die Unterschrift der Partei oder ihrer Vertretung, deren Vollmacht oder die vorgeschriebenen Beilagen oder ist die Vertretung nicht zugelassen, so wird eine angemessene Frist zur Behebung des Mangels angesetzt mit der Androhung, dass die Rechtsschrift sonst unbeachtet bleibt.

⁶ Unleserliche, ungebührliche, unverständliche, übermässig weitschweifige oder nicht in einer Amtssprache verfasste Rechtsschriften können in gleicher Weise zur Änderung zurückgewiesen werden.

⁷ Rechtsschriften, die auf querulatorischer oder rechtsmissbräuchlicher Prozessführung beruhen, sind unzulässig.

Beispiel: Bundesgerichtsgesetz (BGG)

2. Kapitel: Allgemeine Verfahrensbestimmungen

1. Abschnitt: Zuständigkeit

Art. 29 Prüfung

¹ Das Bundesgericht prüft seine Zuständigkeit von Amtes wegen.

² Bestehen Zweifel, ob das Bundesgericht oder eine andere Behörde zuständig ist, so führt das Gericht mit dieser Behörde einen Meinungsaustausch.

Art. 30 Unzuständigkeit

¹ Erachtet sich das Bundesgericht als nicht zuständig, so tritt es auf die Sache nicht ein.

² Hat sich in einem Meinungsaustausch die Zuständigkeit einer anderen Behörde ergeben oder erscheint die Zuständigkeit einer anderen Bundesbehörde als wahrscheinlich, so überweist das Bundesgericht die Sache der betreffenden Behörde.

Art. 31 Vorfragen

Ist das Bundesgericht in der Hauptsache zuständig, so befindet es auch über die Vorfragen.

Art. 42 Rechtsschriften

¹ Rechtsschriften sind in einer Amtssprache abzufassen und haben die Begehren, deren Begründung mit Angabe der Beweismittel und die Unterschrift zu enthalten.

² In der Begründung ist in gedrängter Form darzulegen, inwiefern der angefochtene Akt Recht verletzt. Ist eine Beschwerde nur unter der Voraussetzung zulässig, dass sich eine Rechtsfrage von grundsätzlicher Bedeutung stellt oder ein besonders bedeutender Fall nach Artikel 84 vorliegt, so ist auszuführen, warum die jeweilige Voraussetzung erfüllt ist.

³ Die Urkunden, auf die sich die Partei als Beweismittel beruft, sind beizulegen, soweit die Partei sie in Händen hat; richtet sich die Rechtsschrift gegen einen Entscheid, so ist auch dieser beizulegen.

⁴ Bei elektronischer Zustellung muss das Dokument, das die Rechtsschrift und die Beilagen enthält, von der Partei oder ihrem Vertreter beziehungsweise ihrer Vertreterin mit einer anerkannten elektronischen Signatur versehen werden. Das Bundesgericht bestimmt in einem Reglement, in welchem Format die elektronische Zustellung erfolgen kann.

⁵ Fehlen die Unterschrift der Partei oder ihrer Vertretung, deren Vollmacht oder die vorgeschriebenen Beilagen oder ist die Vertretung nicht zugelassen, so wird eine angemessene Frist zur Behebung des Mangels angesetzt mit der Androhung, dass die Rechtsschrift sonst unbeachtet bleibt.

⁶ Unleserliche, ungebührliche, unverständliche, übermässig weit-schweifige oder nicht in einer Amtssprache verfasste Rechtsschriften können in gleicher Weise zur Änderung zurückgewiesen werden.

⁷ Rechtsschriften, die auf querulatorischer oder rechtsmissbräuchlicher Prozessführung beruhen, sind unzulässig.

Die Eugen-Huber-Regel

Stilrichtlinie

- Pro Artikel höchstens drei Absätze.
- Pro Absatz ein Satz.

Fehlermodellierung

- Artikel mit vier oder mehr Absätzen
- Absätze mit zwei oder mehr Sätzen

Vorverarbeitung

- Artikelgrenzen erkennen
- Absatzgrenzen erkennen
- Satzgrenzen erkennen



Eugen Huber (1849–1923),
Autor des Schweizerischen
Zivilgesetzbuchs (ZGB)

Anmerkung

Benutzer muss einstellen können,
welche Stilregeln überprüft werden.

Das Modalverb *sollen*

Stilrichtlinie

- Das Modalverb *sollen* ist zu vermeiden.
- Ausnahme: Zweckartikel

Fehlermodellierung

- Wortformen von *sollen*
- ... ausser in Zweckartikeln

Vorverarbeitung

- Lemmatisierung
- Zweckartikel erkennen anhand von:
 - **Sachüberschrift** (enthält Begriffe wie *Zweck*)
 - **Position** im Text (am Anfang)

Beispiel: Güterverkehrsverlagerungsgesetz (GVVG)

Art. 1 Zweck

¹ Zum Schutz des Alpengebietes soll der alpenquerende Güterschwerverkehr auf nachhaltige Weise von der Strasse auf die Schiene verlagert werden.

² Zwischen den Verkehrsträgern des alpenquerenden Güterschwerverkehrs soll ein logisch ausgewogenes und den wirtschaftlichen Bedürfnissen entsprechendes Verhältnis bestehen.

Art. 2 Geltungsbereich

Dieses Gesetz gilt für sämtliche Verkehrsträger, soweit sie einen Einfluss auf den alpenquerenden Güterschwerverkehr haben.

Art. 3 Verlagerungsziel

¹ Für den alpenquerenden Güterschwerverkehr auf den Transitstrassen im Alpengebiet (Art. 2 des BG vom 17. Juni 1994⁴ über den Strassentransitverkehr im Alpengebiet) gilt das Ziel von höchstens 650 000 Fahrten pro Jahr.

² Dieses Ziel soll spätestens zwei Jahre nach Inbetriebnahme des Gotthard-Basistunnels erreicht werden.

³ Das Ziel ist auf Dauer einzuhalten und darf nur in einzelnen Jahren mit besonders starker Wirtschafts- und Verkehrsentwicklung überschritten werden.

⁴ Ab dem Jahr 2011 soll das Zwischenziel von höchstens 1 Million Fahrten pro Jahr nicht überschritten werden.

Legaldefinitionen erkennen

Studien im Bereich Legal Information Retrieval

- de Maat & Winkels (2010): maschinelle Erkennung von **Legaldefinitionen in niederländischen Gesetzestexten**
- Walter & Pinkal (2009): maschinelle Erkennung von **Definitionen in deutschen Gerichtsentscheiden**

Ansatz

- Legaldefinitionen anhand von **typischen Satzmustern** erkennen

Resultate

- 100% Präzision für niederländische Gesetzestexte
- 70% Präzision für deutsche Gerichtsentscheide

Quellen

de Maat, E., Winkels, R. (2010) Automated classification of norms in sources of law. In: *Semantic Processing of Legal Texts*. Berlin, Springer.

Walter, S., Pinkal, M. (2009) Definitions in court decisions: Automatic extraction and ontology acquisition. In: *Law, Ontologies and the Semantic Web*. Amsterdam, IOS Press.

Satzmuster von Legaldefinitionen

Beispiel

Als **Geflügel** im Sinne dieser Verordnung gelten **Hühnervögel (Galliformes), Schwimmvögel (Anseriformes) und Laufvögel (Struthioniformes)**.

Satzmuster: Als **NP** im Sinne dieser Verordnung gilt/gelten **NP**.
Definiendum *Definiens*

Beispiel

Gehege: **umgrenzter Bereich, in dem Tiere gehalten werden, einschliesslich Auslaufflächen, Käfigen, Volieren, Terrarien, Aquarien, Aufzuchtbecken und Fischeichen**.

Satzmuster: **NP** : **NP** [, einschliesslich/insbesondere/namentlich **NP**].

Vorverarbeitung: Chunking

Quelle

Bratschi, R. (2009) «Frau im Sinne dieser Badeordnung ist auch der Bademeister». Legaldefinitionen aus redaktioneller Sicht. *LeGes*, 20(2):191–213.

Unnötige Legaldefinitionen

Stilrichtlinie

- Legaldefinition ist nur nötig, wenn der Begriff mind. dreimal vorkommt

Fehlermodellierung

- Definiendum einer Legaldefinition, das im restlichen Text weniger als dreimal vorkommt

Zirkeldefinitionen

Stilrichtlinie

- Begriff nicht mit sich selber definieren

Fehlermodellierung

- Legaldefinition, in der das Definiendum (oder Elemente davon) auch im Definiens vorkommen

Redaktionsbeispiele

Früchte im Sinne dieser Verordnung sind:

- a. gewaschene und gerüstete Früchte;
- b. [...]

Als Küchen- und Speisereste gelten Speisereste, die aus Einrichtungen stammen, in denen Lebensmittel für den unmittelbaren Verzehr hergestellt werden, wie Restaurants, Catering-Einrichtungen und Küchen, einschliesslich Gross- und Haushalbküchen.

¹ Als «private Hausangestellte» gelten [...] die im häuslichen Dienst [...] beschäftigten Personen.

² [...]

³ Der Begriff «häuslicher Dienst» bezeichnet alle von privaten Hausangestellten wahrgenommenen Tätigkeiten [...].

Die Eugen-Huber-Regel

Stilrichtlinie

- Pro Artikel höchstens drei Absätze.
- Pro Absatz ein Satz.
- Pro Satz eine Aussage.



Eugen Huber (1849–1923),
Autor des Schweizerischen
Zivilgesetzbuchs (ZGB)

Problem

Diese Regel bezieht sich nicht mehr nur auf die **Form**, sondern auch auf die **Bedeutung** des Textes.

Lösung / Forschungsfrage

Wir benötigen **Indikatoren** («Symptome») in der **Form** des Textes, die auf das Vorhandensein einer solchen Richtlinienverletzung hinweisen.

Beispiel: weiterführende Relativsätze

³ Die berufliche Vorsorge wird durch die Beiträge der Versicherten finanziert, **wobei** die Arbeitgeberinnen und Arbeitgeber mindestens die Hälfte der Beiträge ihrer Arbeitnehmerinnen und Arbeitnehmer bezahlen.

besser

³ Die berufliche Vorsorge wird durch die Beiträge der Versicherten finanziert.
^{3bis} Die Arbeitgeberinnen und Arbeitgeber bezahlen mindestens die Hälfte der Beiträge ihrer Arbeitnehmerinnen und Arbeitnehmer.

Quelle

Höfler, S. (2011) «Ein Satz – eine Aussage». Multipropositionale Rechtssätze an der Sprache erkennen. *LeGes*, 22(2):259–280.

Beispiel: als Adverbiale formulierte Vorbehalte

Vorbehältlich einer anderen Abmachung werden für die ausgezeichneten Filme die Preise je zur Hälfte an die Produktion und an die Regie ausbezahlt.

besser

¹ Für die ausgezeichneten Filme werden die Preise je zur Hälfte an die Produktion und an die Regie ausbezahlt.

² Anderslautende Abmachungen bleiben vorbehalten.

Quelle

Höfler, S. (2011) «Ein Satz – eine Aussage». Multipropositionale Rechtssätze an der Sprache erkennen. *LeGes*, 22(2):259–280.

Beispiel: in Aufzählungen eingeschobene Sätze

Das Amt:

- a. prüft Gesuche; diese müssen schriftlich und begründet eingereicht werden;
- b. erteilt die Bewilligung, wenn die Voraussetzungen erfüllt sind.

besser

¹ Das Amt:

- a. prüft Gesuche;
- b. erteilt die Bewilligung, wenn die Voraussetzungen erfüllt sind.

² Die Gesuche sind zu begründen und schriftlich einzureichen.

Quelle

Höfler, S. (2011) «Ein Satz – eine Aussage». Multipropositionale Rechtssätze an der Sprache erkennen. *LeGes*, 22(2):259–280.

Geltungsbereichsbestimmungen

Stilregel

Geltungsbereichsbestimmungen sollen **keine materiellen Teile** enthalten.

Beispiel

Diese Verordnung gilt nicht für:

- a. technisches Personal;
- b. Hilfsassistierende; stellt ein Institut solche Personen ein, **muss** das Dekanat unterrichtet werden;
- c. Mitarbeitende von Sekretariaten und Verwaltung.

Indikatoren

für die Präsenz eines materiellen Elements:

- Modalverb ***muss***
- Abweichung vom typischen Satzmuster für Geltungsbereichsbestimmungen

Grenzen des Ansatzes

Stilregeln

- Ein Begriff sollte nur definiert werden, wenn er ohne Definition **missverständlich, unverständlich** oder **strittig** ist.
- Ein definierter Begriff ist innerhalb eines Erlasses immer **im definierten Sinn** zu verwenden.
- Die Definition soll **zweckmässig** sein im Hinblick auf den Regelungszweck.
- Die Definition soll **(in einem vernünftigen Mass) präzise** sein.
- Die Definition soll **adressatengerecht** formuliert sein und so weit wie möglich am allgemeinen Sprachgebrauch anknüpfen.

Zusammenfassung

Ziel

- Weiterentwicklung der **Methoden der maschinellen Stilprüfung** in der technischen Redaktion für den Einsatz **in der Gesetzesredaktion**

Herausforderungen

- domänenspezifische **Vorverarbeitung**
- domänenspezifische **Fehlermodellierung**

Zusammenfassung

Forschungsfragen

- Für jede Richtlinie: Gibt es **sprachliche Indikatoren** für eine Verletzung?
- Wie und wie zuverlässig kann man diese **maschinell erkennen**?

Machbarkeit

Je nach Abstraktionsgrad einer Richtlinie ist eine maschinelle Überprüfung:

- einfach zu realisieren
- aufwändiger, aber grundsätzlich machbar
- machbar, aber mit eingeschränkter Genauigkeit und Trefferquote
- nicht möglich

Cui bono?

Gesetzesredaktoren / Gesetzesredaktorinnen

Vorteil

Auch wenn bei weitem nicht alle Richtlinien maschinell überprüft werden können, so wird doch ein Grundstock an potentiellen Richtlinienverletzungen eruiert, die nicht mehr mühsam «von Hand» gesucht werden müssen.

- ➡ spart Zeit und Aufwand

Autoren / Autorinnen von Erlassen

Vorteil

Man muss nicht alle Richtlinien auswendig kennen, sondern wird vom System von Fall zu Fall auf die entsprechende Richtlinie hingewiesen.

- ➡ Richtlinien würden evtl. konsequenter angewendet

Beteiligte Personen

Mitarbeitende

- Dr. **Stefan Höfler**
- lic. phil. **Alexandra Bünzli**
- **Kyoko Sugisaki**, M.A.
- Dr. **Cathrin Senn**

Wissenschaftliche Begleitung

- Prof. Dr. **Michael Hess**, Institut für Computerlinguistik, UZH
- Prof. Dr. **Felix Uhlmann**, Rechtswissenschaftliches Institut, UZH
- Dr. **Rebekka Bratschi**, Zentrale Sprachdienste, Bundeskanzlei