

Klassifikation von Figurenauf- und -abtritten in XML-kodierten Dramen

Motivierendes Textbeispiel

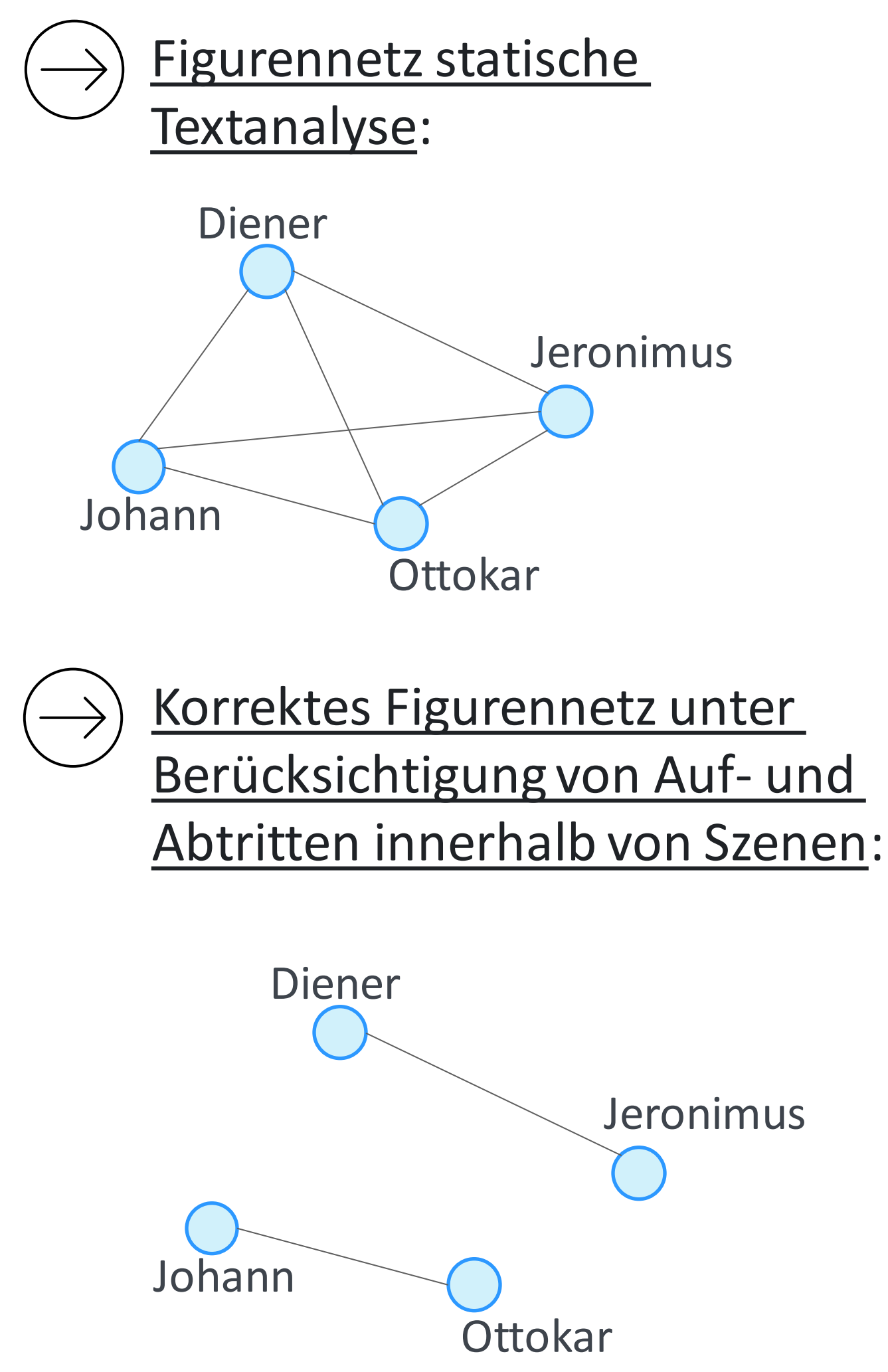
EIN DIENER
tritt auf.
War nicht Graf Rupert hier?

JERONIMUS.
Suchst du ihn? Ich geh mit dir.

Alle ab.
Ottokar und Johann treten von
der andern Seite auf.

OTTOKAR.
Wie kamst du denn zu diesem Schleier?

(Kleist, Familie Schroffenstein,
Erster Aufzug, erste Szene)



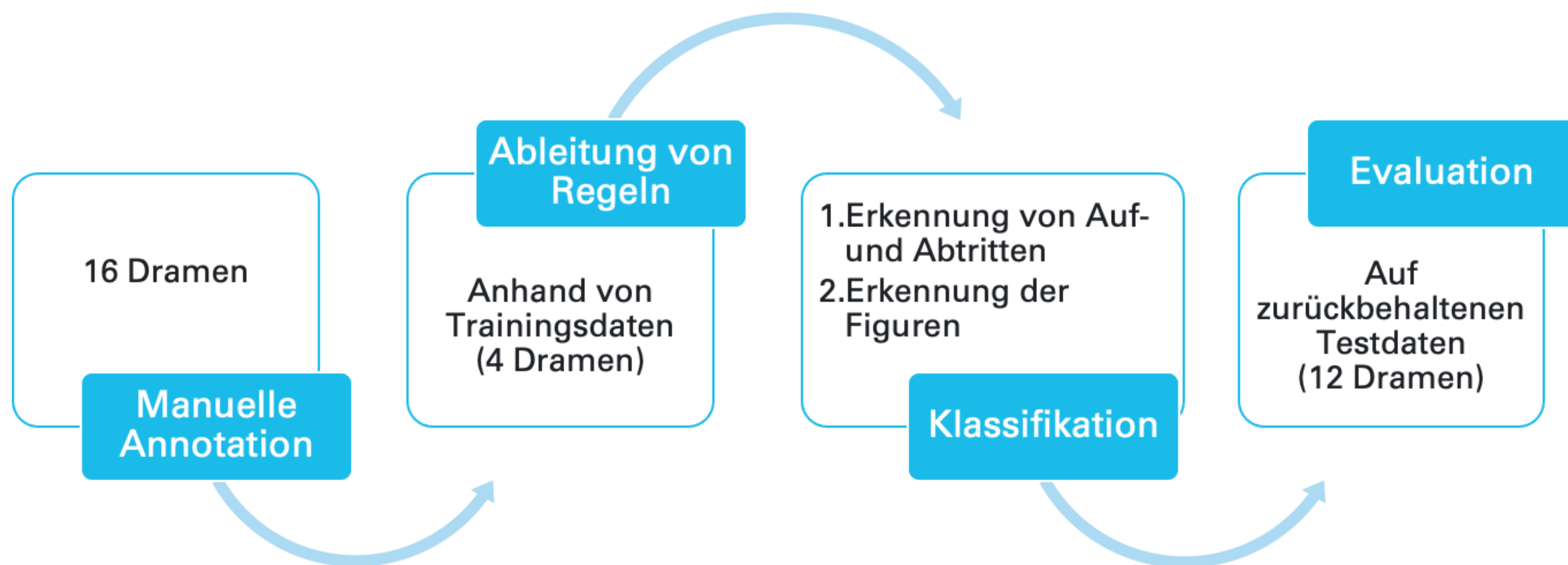
Datengrundlage und Ausgangssituation

- Datengrundlage: 16 Dramentexte aus GerDraCor
- TEI-kodiert
 - Regieanweisungen sind als *stage*-Elemente gekennzeichnet.
 - Bisher enthalten die Tags der Regieanweisungen keine zusätzlichen Attribute über deren Inhalt.

```
<sp who="#wagner">
  <speaker>WAGNER.</speaker>
  <lg>
    <l>Ich hätte gern nur immer fortgewacht,</l>
    <l>Um so gelehrt mit Euch mich zu besprechen.</l>
    <l>Doch morgen, als am ersten Ostertage,</l>
    <l>Erlaubt mir ein' und andre Frage.</l>
    <l>Mit Eifer hab' ich mich der Studien beflissen;</l>
    <l>Zwar weiß ich viel, doch möcht' ich alles wissen.</l>
  </lg>
  <stage>Ab.</stage>
</sp>
```

→ `<stage type="exit" who="#wagner">Ab.</stage>`

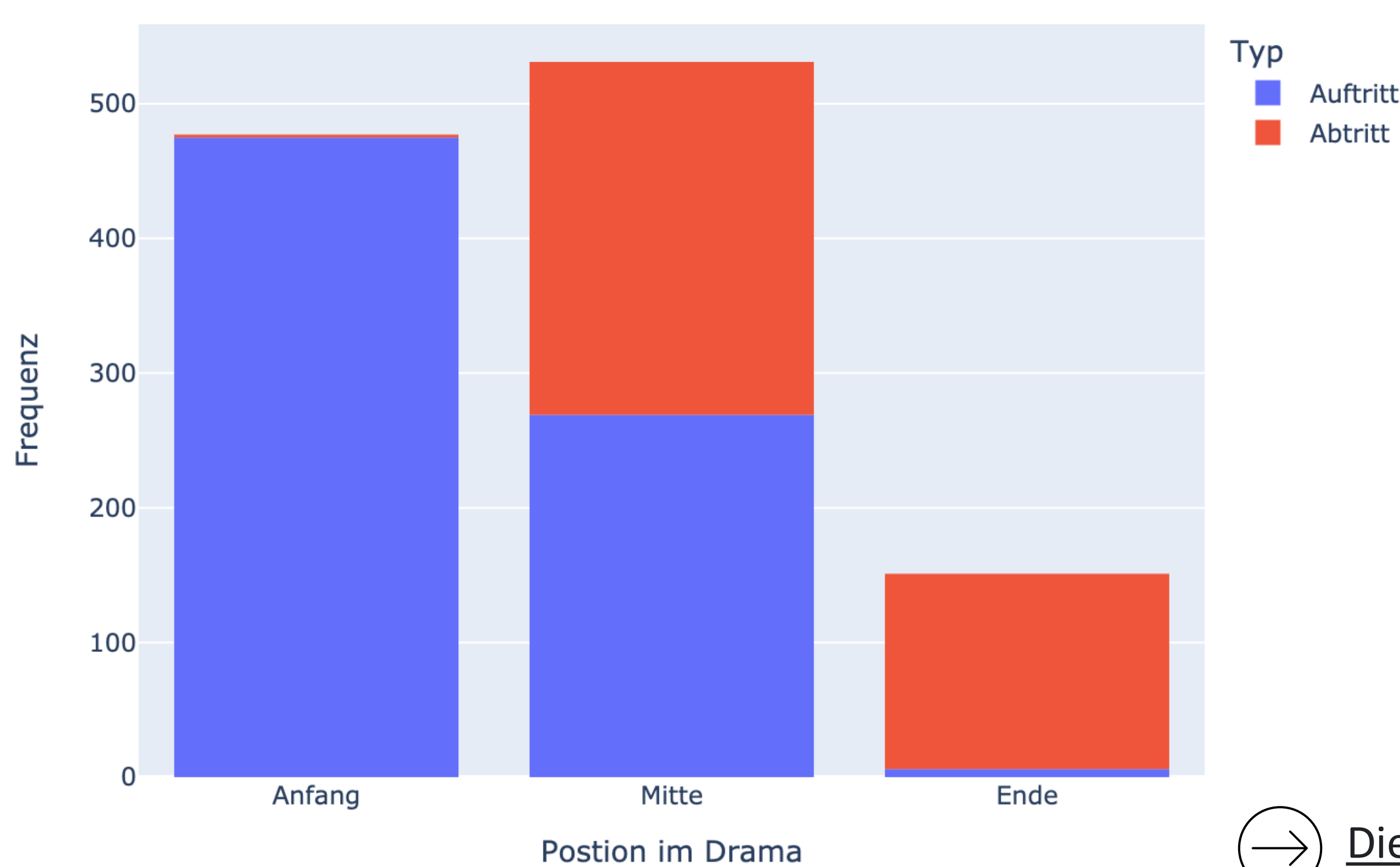
Semi-automatischer Workflow



In den Regeln verwendete Features:

- 🔑 Schlüsselwörter und -phrasen als reguläre Ausdrücke
- 📍 Position der Regieanweisung innerhalb einer Szene
- 👤 Figurenname oder zugeordnete Figurenrede

Analyse & Evaluation



	Auf-/Abtritterkennung	Figurenerkennung
Precision	0,85 ± 0,17	0,93 ± 0,08
Recall	0,85 ± 0,17	0,81 ± 0,18
F ₁	0,85 ± 0,16	0,87 ± 0,14

- Wichtigste Fehlerquellen:**
- Von den aufgestellten Regeln stark abweichende Texte
 - Überdurchschnittlich lange Regieanweisungen
 - Von der Figurenliste abweichende Namen

→ Die meisten Auf- und Abtritte (46%) finden innerhalb von Szenen statt.

Alle Daten und Skripte zu diesem Beitrag sind hier verfügbar:

