

# Annotation und Analyse narrativen Raums in lateinamerikanischen Romanen des 19. Jahrhunderts

*Ulrike Henny-Krahmer, Universität Rostock*

*Vortrag im Kolloquium »Phänomenologie der Digital Humanities«, Wintersemester 2023/24 an der FU Berlin, 30.11.2023*

## Überblick

1. Kontext: Projekt CANSpiN
2. Forschungsfragen und Korpus: Raum im lateinamerikanischen Roman des 19. Jahrhunderts
3. Annotation raumreferentieller Ausdrücke



# 1. Kontext: Projekt CANSpiN

# 1. Kontext

## Projekt CANSpiN

- DFG-Projekt Computational Approaches to Narrative Space in 19th and 20th Century Novels (2023-2026), <https://www.canspin.uni-rostock.de/>
- Teilprojekt im Schwerpunktprogramm „Computational Literary Studies“ (SPP 2207), <https://dfg-spp-cls.github.io/>
- Ziel: Entwicklung von computergestützten Methoden zur Erkennung räumlicher Entitäten in narrativen Texten

## 1. Kontext

### Projekt CANSpiN

- **Laufzeit:** April 2023 bis März 2026, an der Universität Rostock
- **Projektleitung:** Holger Helbig (Germanistik), Ulrike Henny-Krahmer (Digital Humanities), Roger Labahn (Mathematik)
- **Mitarbeiter:** Marc Lemke (DH), Nils Kellner (Doktorandenstelle DH)
- **Studentische Hilfskräfte:** Nikita Vasyiltsov (Germanistik), Jasko Miebs (Mathematik), Nils Schröder (Mathematik)
- **Weitere Projektgruppe:** Astrid Köhler (Queen Mary University of London), Gimena del Rio Riande (CONICET, Buenos Aires), Tobias Strauß (Universität Rostock, Mathematik), Konrad Sperfeld (Planet AI GmbH, Rostock), Jochen Zöllner (Planet AI GmbH, Rostock)

## 1. Kontext

### Projekt CANSpiN: Ausgangspunkt

- **narrativer Raum**, hier: konkreter Raum der erzählten Welt, in dem Figuren leben, handeln, sich bewegen
- **Literaturtheorie**: ausgearbeitete Konzepte zum narrativen Raum (u.a. Dennerlein 2009; Piatti 2008; Ryan 2014; Lotman 1977; Renner 2004)
- **Computational Literary Studies**: einige Ansätze zur Formalisierung von Raumkonzepten und Abbildung auf Textmerkmale (u.a. Viehhauser und Barth 2017; Barth 2021, 2022; Viehhauser 2020; Schumacher 2022, 2023)

## 1. Kontext

### Projekt CANSpiN: Ausgangspunkt

Um halb ein Uhr hielt [der Zug]<sub>setting</sub> [auf der Station Benares]<sub>setting</sub>.

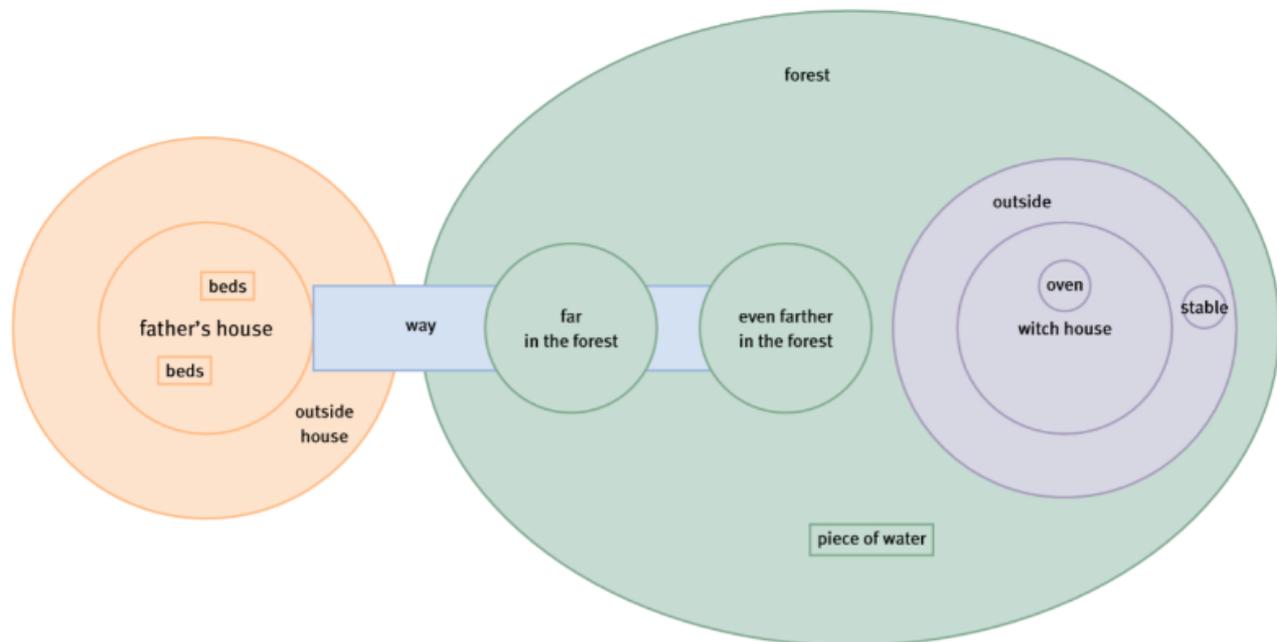
[Das nicht eben prachtvolle Haus [in Savile Row]<sub>mention</sub>]<sub>mention</sub> empfahl sich durch größte Bequemlichkeit.

(Beispiele für Raumentitäten und ihre Handlungsrelevanz aus Barth (2021), aus Jules Verne "Reise um die Erde in 80 Tagen")

## 1. Kontext

### Projekt CANSpiN: Ausgangspunkt

Analyse der  
*spatial frames*  
nach Viehhauser  
(2020), zu der  
englischen  
Übersetzung vom  
grimm'schen  
Märchen „Hänsel  
und Gretel“



## 1. Kontext

### Projekt CANSpiN: Ziele

- Entwicklung von **computergestützten Methoden zur Erkennung** räumlicher Entitäten
  - › für verschiedene Arten von Verweisen auf den narrativen Raum in literarischen Texten
  - › z.B. Toponyme ("Rostock", "Göttingen"), allgemeine raumbezogene Substantive ("Zimmer", "Kirche", "Straße", "Berg"), deiktische Ausdrücke ("hier", "dort")
  - › unter Verwendung bestehender semantischer Wortnetze und Techniken des maschinellen Lernens, insb. neuronaler Netze

## 1. Kontext

### Projekt CANSpiN: Ziele

- Entwicklung von Verfahren zur Zuordnung räumlicher Entitäten zu **narrativen Funktionen**
  - } z.B. handlungsrelevante Erwähnungen von Raum vs. Beschreibungen des räumlichen Settings vs. weitere Referenzen auf Raum
- Annäherung an die **semantische Struktur** von Erzählwelten und die symbolische Bedeutung von Teilräumen
  - } z.B. raumsemantische Struktur eines Textes nach Lotman (1977)
  - } z.B. Verbindung der Erkennung von Raumentitäten mit Topic Modeling und Sentimentanalyse

## 1. Kontext

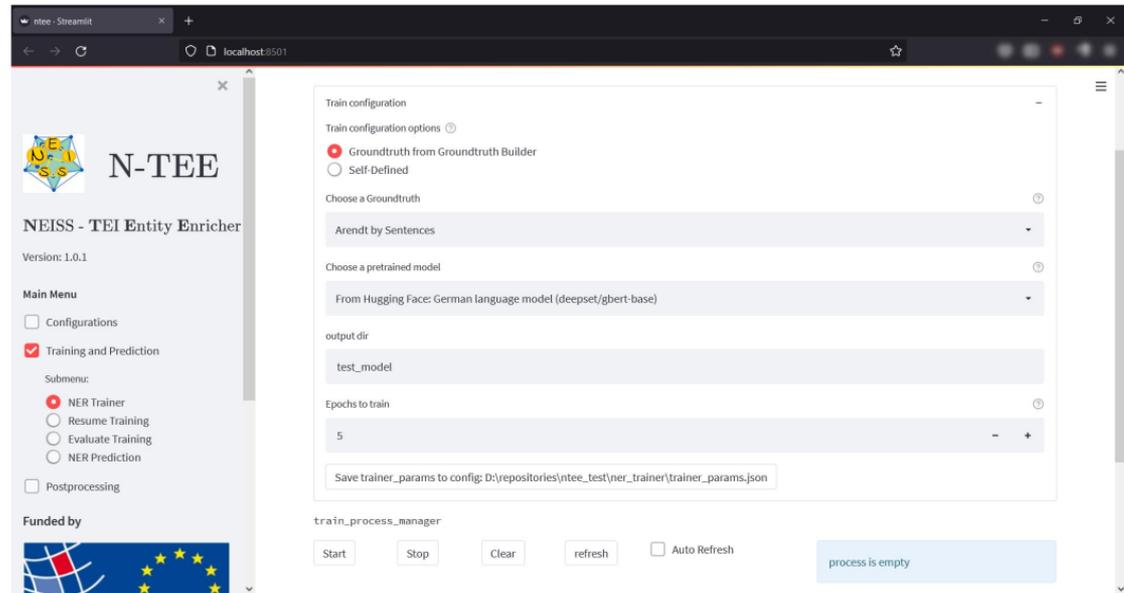
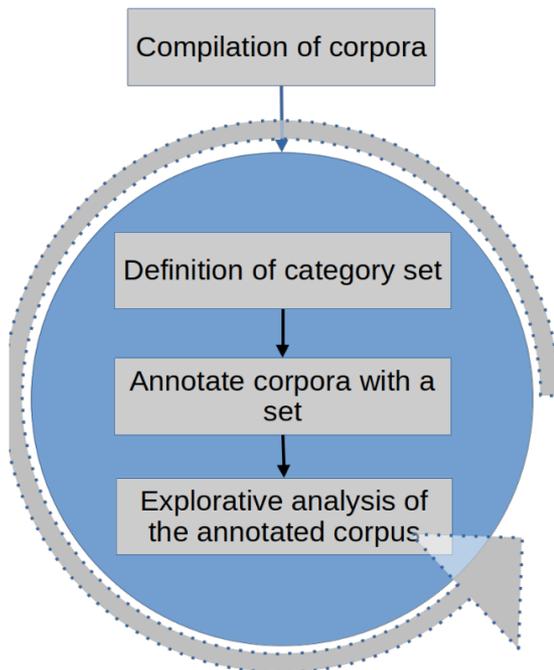
### Projekt CANSpiN: Korpora

- **deutschsprachige Romane** aus dem **19. Jahrhundert** (1850-1910)
  - } Ziel: ca. 200 Romane; Quellen: ELTeC-deu (100 Romane), DTA, Text Grid Repository
- **spanischsprachige Romane** aus dem **19. Jahrhundert** (1850-1910)
  - } Ziel: ca. 200 Romane; Quellen: Corpus de novelas hispanoamericanas del siglo XIX (Conha19), Corpus of Novels of the Spanish Silver Age (CoNSSA)
- **deutschsprachige Romane** aus dem **20. Jahrhundert** (1950-2000)
  - } Ziel: ca. 50 Romane; Quellen: TEI-Dateien Uwe Johnson-Werkausgabe, E-Books
  - } Autor\*innen: Heinrich Böll, Christine Brückner, Uwe Johnson, Walter Kempowski, Christa Wolf, ...

# 1. Kontext

## Projekt CANSpiN: Methoden

NEISS TEI Entity Enricher (NTEE):  
[https://github.com/NEISSproject/tei\\_entity\\_enricher](https://github.com/NEISSproject/tei_entity_enricher)





## 2. Forschungsfragen und Korpus: Raum im lateinamerikanischen Roman des 19. Jahrhunderts

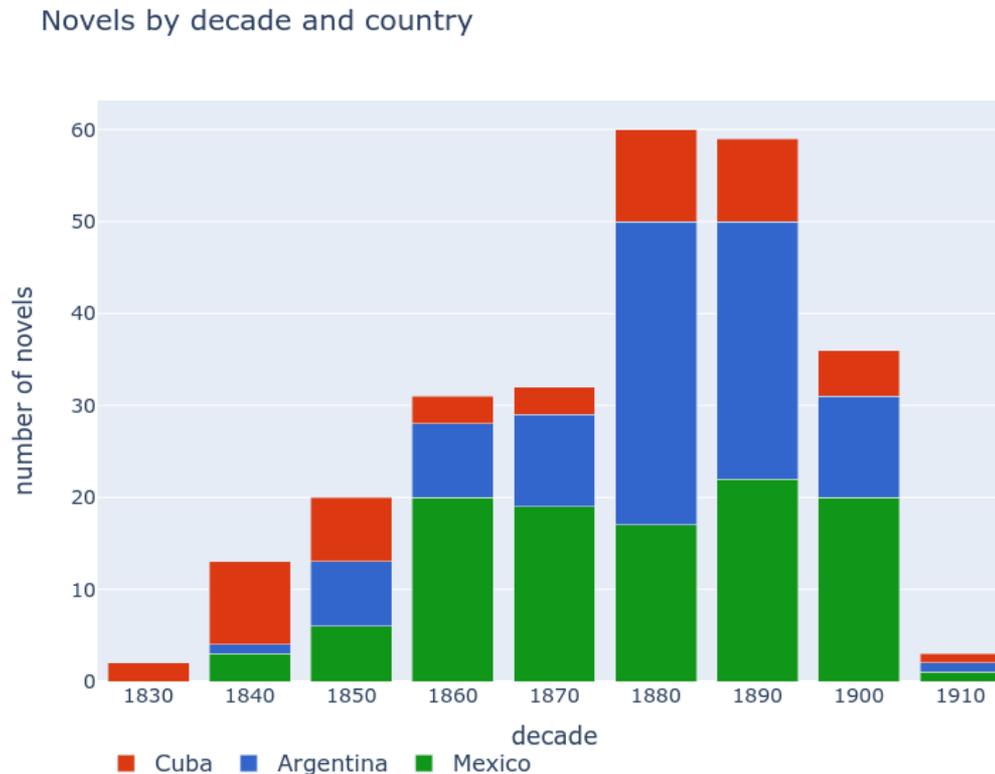
## 2. Forschungsfragen und Korpus

### Raum im lateinamerikanischen Roman des 19. Jahrhunderts

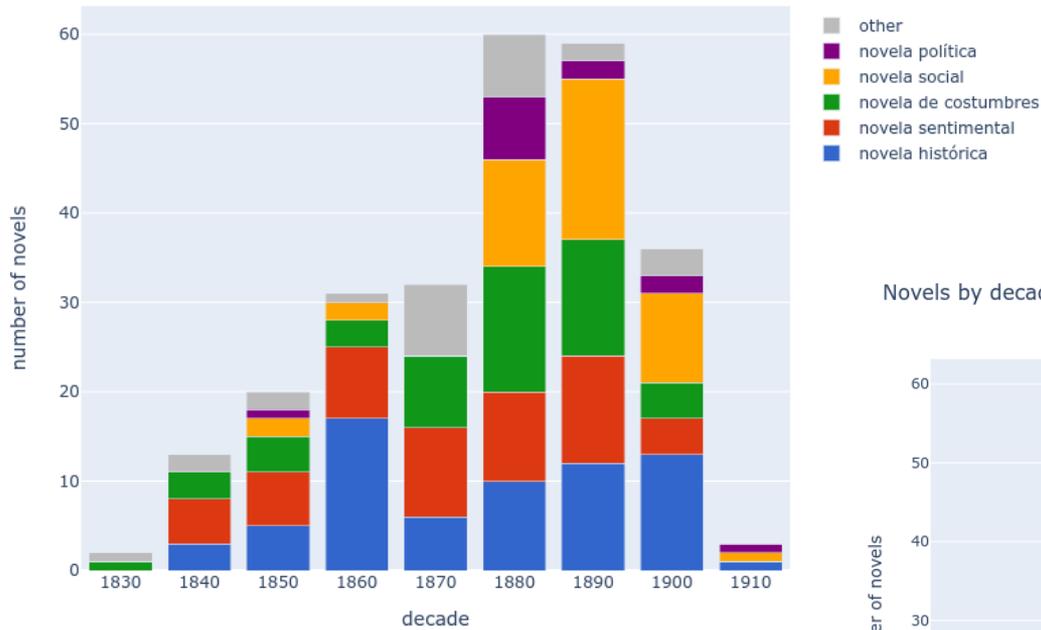
- literaturgeschichtliche Fragen:
  - } Untersuchung des Zusammenhangs von Darstellung von Raum in den Romanen mit Fragen **nationaler, regionaler und anderer raumbezogener Identitäten**, in Anlehnung an das Konzept der "imagined communities" (Anderson 1983)
  - } für lateinamerikanische Romane insb. Bezüge zwischen Romantexten und dem Prozess der **Nationenbildung im 19. Jahrhundert**
  - } z.B. Verknüpfung der Analyse der räumlichen Bezüge in den Texten mit den Nationalitäten der Autoren, den Erscheinungsorten der Romane, dem Setting und den Themen der Texte
  - } auch: Zusammenhang von Raum und **Untergattungen** des Romans

## Ausgangspunkt: Korpus Conha19

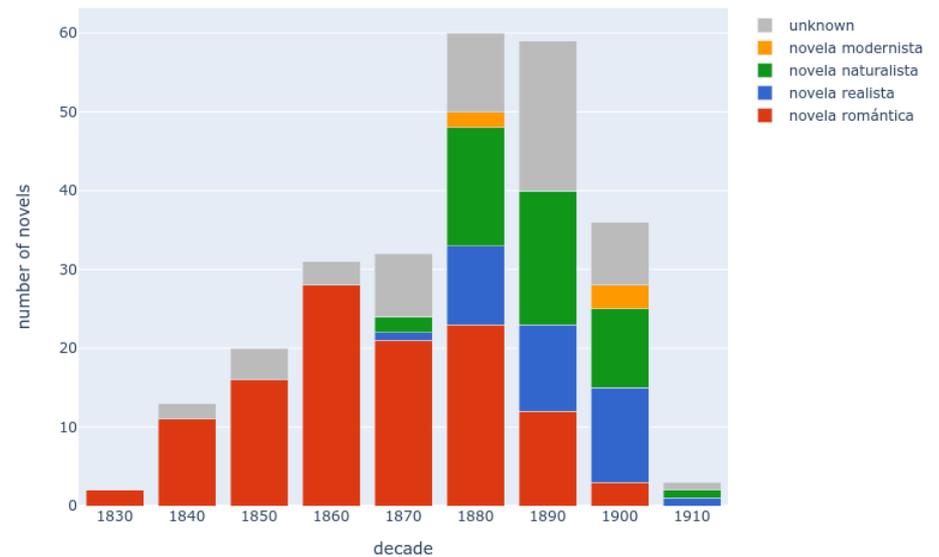
<https://github.com/cligs/conha19>



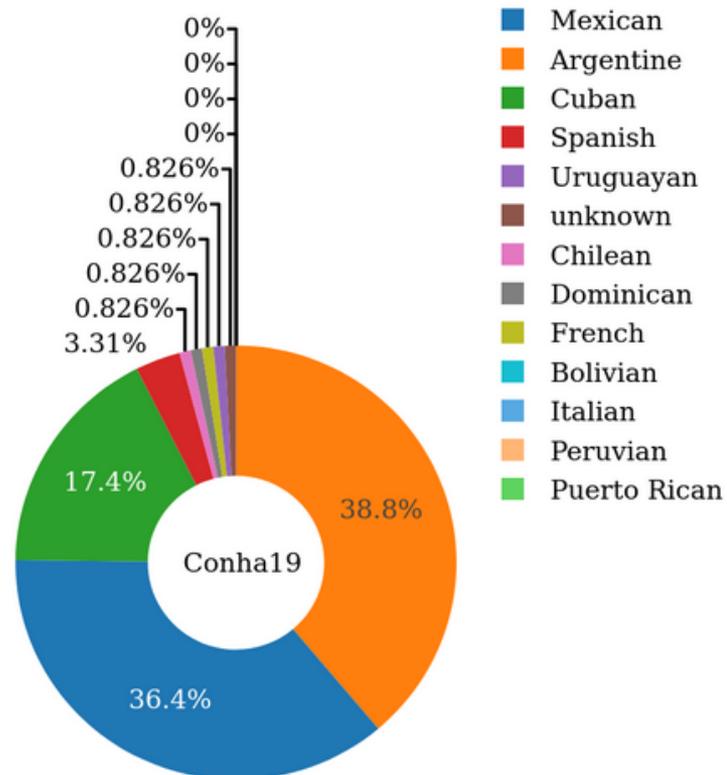
Novels by decade and thematic subgenre



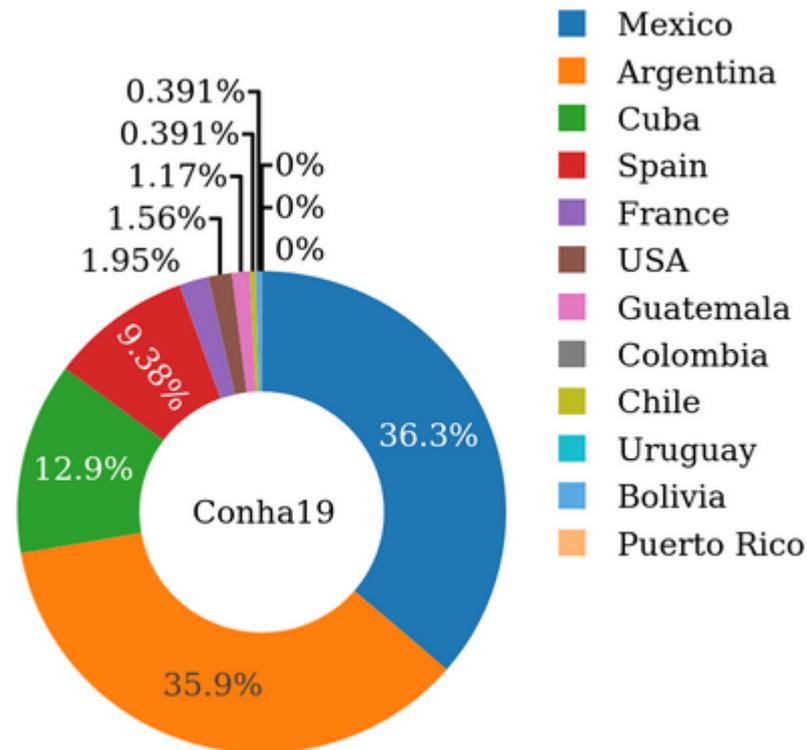
Novels by decade and literary current



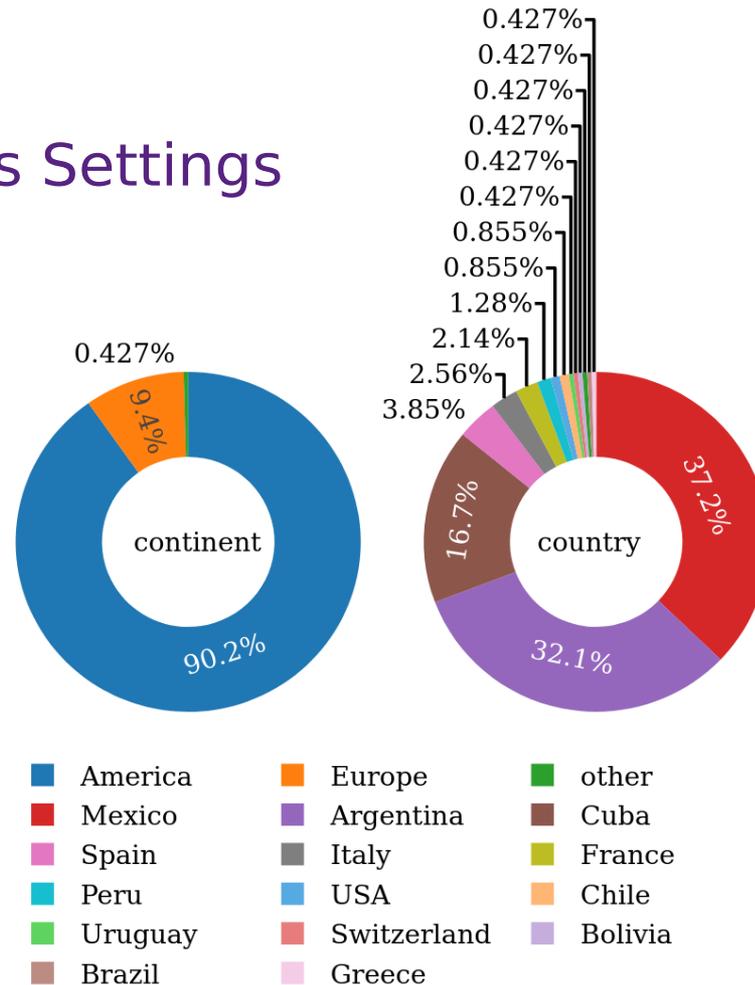
## Autor:innen nach Nationalität



## Publikationsorte der Erstausgaben



## Kontinent und Land des Settings





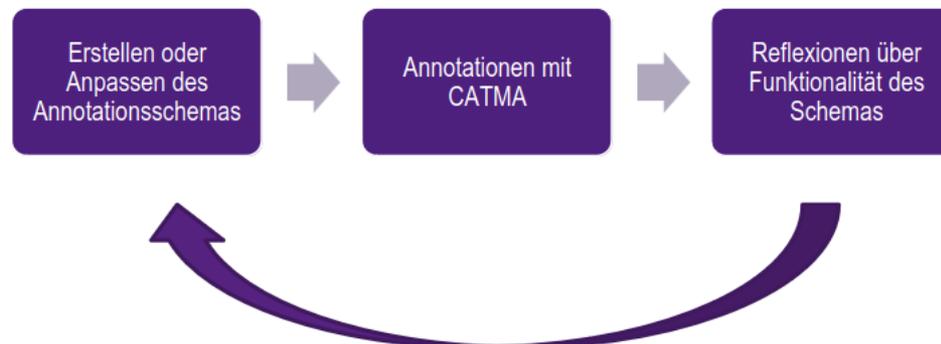
### 3. Annotation raumreferentieller Ausdrücke

### 3. Annotation raumreferentieller Ausdrücke

- offenes, exploratives Verfahren: Erproben verschiedener raumanalytischer **Kategorien-Sets**
  - computergestützte Operationalisierbarkeit?
  - literaturwissenschaftliche Aussagekraft?
- Entwicklung von **Annotationsrichtlinien** für jedes Kategorienset
  - aktuell: **CANSpiN.CS1**
    - Erfassen raumreferentieller Ausdrücke zunächst ohne Berücksichtigung narrativer Funktionen der Ausdrücke und ihres Verhältnisses zur Diegese
    - Erfassen der Räumlichkeit der Texte an der Textoberfläche
- aktuell: Annotationen mit dem Tool CATMA (<https://catma.de/>)

## Annotationen: Relevante Vorarbeiten

- fuzzy-set von Raumindikatoren zur Analyse von Räumlichkeit in Romanen von **Schumacher (2023)**
- Systematik raumreferentieller Ausdrücke zur Bezeichnung des konkreten Raumes der erzählten Welt von **Dennerlein (2009)**
- Empfehlungen zur Entwicklung von Annotationsrichtlinien von **Reiter (2020)**



## Kategorien in CANSpiN.CS1

- Kategorien:
  - } **Ort**
  - } **Richtung**
  - } **Positionierung**
  - } **Dimensionierung**
  - } **Bewegung**
- jeweils unterschieden: übertragender und nicht übertragender Gebrauch

## Kategorien in CANSpiN.CS1: Ort

- **Orte** sind „alles, was als fester, unbeweglicher Punkt oder als feste, unbewegliche Größe beliebiger Ausdehnung [...] identifiziert werden kann.“ (Schumacher 2023: 64)
- in CANSpiN außerdem:
  - } auch bewegte Größen, sofern sie das räumliche Referenzsystem darstellen (z.B. das Innere von Fahrzeugen)
  - } prinzipiell bewegliche Entitäten können als Orte fungieren, insofern sie in ihrem Referenzsystem in der dargestellten Situation nicht bewegt und als Orte angesprochen werden, etwa als Ausgangs- und Zielpunkt von Bewegungen (z.B. Möbelstücke)
- **Indikatoren:** Orte sind Nomen, etwa Toponymika, Eigennamen und Gattungsbezeichnungen (vgl. Dennerlein 2009: 209)

## Kategorien in CANSpiN.CS1: Beispiele Ort

- (1) Sie stand im **Zimmer**.
- (2) Sie wohnten in **New York City**.
- (3) Ich zog in die **Stadt**.
- (4) Sie schwebten oberhalb des **Hauses**.
- (5) Er stellt sich vor, auf die **Bergspitze** zu gehen.
- (6) Durch die **Gipfel** fällt ein Licht auf die **Bäume**.
- (7) Er blickte in den **Himmel**.
- (8) Die Professorin sprang von ihrem **Arbeitstisch** auf.

## Kategorien in CANSpiN.CS1: Richtung

(19) Sie bogen nach Süden ab.

(20) Er ging hoch nach oben auf die höchste Bergspitze.

(21) Ich zog weiter südlich in die Stadt.

(22) Sie kam aus dem oberen Teil des Hauses.

- **Richtungen** sind subjektabhängige und -unabhängige Ausdrücke, die im Zusammenhang mit einem Prädikat (häufig Bewegungen), einer attributiv verwendeten Verbform oder als Teil von Nominalphrasen eine Richtungsangabe konstituieren
- **Indikatoren:** Richtungen werden von Präpositionen, Lokal-Adverbien, adverbial verwendeten Adjektiven und richtungsangebenden Nomen ohne Artikel angezeigt

## Kategorien in CANSpiN.CS1: Positionierung

(23) Sie stand **in** dem **linken** Zimmer.

(24) Sie wohnten **im westlichen** Teil **von** New York City.

(25) Es ist später Abend **im** Wald **unserer** Stadt.

(26) Sie schwebten **draußen** **oberhalb** des Hauses.

(27) Das Zelt-Lager A lag 250 Meter **höher** **am** Berg als Lager B.

(28) Sie suchten einen **höhergelegenen** Aussichtspunkt.

- **Positionierungen** kennzeichnen die räumliche Relation zwischen zwei Konstituenten
- **Indikatoren:** Positionierungen sind Präpositionen, Lokal-Adverbien, adverbial und attributiv verwendete Adjektive sowie Possessivpronomen und Artikel

## Kategorien in CANSpiN.CS1: Dimensionierung

(29) Es war ein **großes** Zimmer.

(30) Sie war von **schmäler** Gestalt.

(31) Dieser Kirchturm war viel **höher**.

(32) Durch die **Masse von** Bäumen konnten sie keinen Horizont sehen.

(33) Und **kaum eine Tagreise entfernt** erblühen Universitäten.

(34) Das sind **höchst** unerfreuliche Nachrichten.

(35) Der Weg betrug **250 Meter**.

- **Dimensionierungen** sind im Kern Adjektive, die durch Quantifizierung die räumlichen Eigenschaften von Entitäten oder Abstände zwischen ihnen bestimmen
- **Indikatoren:** Dimensionierungen sind im Kern attributiv und adverbial verwendete Adjektive, Adverbien und Mengenangaben

## Kategorien in CANSpiN.CS1: Bewegung

(36) Er ging in die Natur.

(37) Bei der Kälte der Nacht begann sie zu zittern.

(38) Sie riefen lauthals nach Hilfe.

(39) Durch die Gipfel fällt bleiches Licht auf den Waldweg.

(40) Der Wald lichtet sich.

- **Bewegungen:** Positionen, Dimensionen oder Richtungen von Entitäten im Raum können sich verändern
- **Indikatoren:** Bewegungen sind Verben, Partizip-Konstruktionen und mitunter mehrgliedrige idiomatische Ausdrücke, die ein räumliches Verb enthalten

## Annotationsrichtlinie CANSpiN.CS1

- Ziele:
  - } Einsatz für computergestützte **multi-class-Annotationen**
  - } zur Erzeugung von Trainingsdaten für Deep Learning-Verfahren
- daher:
  - } jedes Token eines Textes soll also ausschließlich **einer Kategorie** zugeordnet werden
  - } **syntagmatischer Referenzrahmen** zur Feststellung der Kategorie von Token: näherungsweise der Satz mit seinem unmittelbaren Kontext

## Fazit

- **Positiv:**

- narrativer Raum literaturtheoretisch schon sehr gut erforscht
- Machine-Learning-Verfahren als vielversprechender Ansatz zur Erkennung narrativen Raums

- **Trotzdem:**

- computergestützte Verarbeitung erfordert Anpassung der Kategorien zur Untersuchung narrativen Raums
- aktuell noch ein weiter Weg von der Erfassung raumreferentieller Ausdrücke an der Sprachoberfläche hin zur computergestützten Erkennung der Struktur von narrativem Raum der erzählten Welt

## Literaturhinweise

- Anderson, Benedict (1983): *Imagined Communities. Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Barth, Florian (2021). "Konzept und Klassifikation literarischer Raumentitäten." In: INFORMATIK 2020, hrsg. von Ralf Heinrich Reussner, Anne Koziolk und Robert Heinrich, 1281–1293. Bonn: Gesellschaft für Informatik.  
[https://doi.org/10.18420/inf2020\\_120](https://doi.org/10.18420/inf2020_120).
- Barth, Florian (2022). "Von der literaturwissenschaftlichen Theorie zu maschinellen Erkennung: Operationalisierung von Raumentitäten und Settings." In: *Digitale Verfahren in der Literaturwissenschaft*, hrsg. von Jan Horstmann and Frank Fischer. Sonderband #6 von Textpraxis. Digitales Journal für Philologie. <https://doi.org/10.17879/64059429732>.
- Dennerlein, Katrin (2009): *Narratologie des Raumes*. Berlin, New York: De Gruyter.  
<https://doi.org/10.1515/9783110219920>.
- Gius, Evelyn, Jan Christoph Meister, Malte Meister, Marco Petris, Mareike Schumacher, Dominik Gerstorfer (2023): *CATMA 7 (Version 7.0)*. Zenodo.  
<https://doi.org/10.5281/zenodo.1470118>.
- Lotman, Jurij Michailowitsch (1977): *Die Struktur literarischer Texte*. München: Fink.

## Literaturhinweise

- Piatti, Barbara (2008): *Die Geographie der Literatur. Schauplätze, Handlungsräume, Raumphantasien*. Göttingen: Wallstein.
- Reiter, Nils (2020): „Anleitung zur Erstellung von Annotationsrichtlinien.“ In: *Reflektierte algorithmische Textanalyse: Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt*. Hrsg. von Nils Reiter, Axel Pichler und Jonas Kuhn. Berlin, Boston: De Gruyter. S. 193–202.  
<https://doi.org/10.1515/9783110693973-009>.
- Renner, Karl Nikolaus (2004): “Grenze und Ereignis. Weiterführende Überlegungen zum Ereigniskonzept von Jurij M. Lotman.” In: *Norm – Grenze – Abweichung: Kultursemiotische Studien zu Literatur, Medien und Wirtschaft*, hrsg. von Gustav Frank und Wolfgang Lukas, S. 357–81. Passau: Verlag Karl Stutz.  
[https://web.archive.org/web/20230717115613/https://www.kultursemiotik.com/wp-content/uploads/2015/01/Renner\\_Grenze-und-Ereignis.pdf](https://web.archive.org/web/20230717115613/https://www.kultursemiotik.com/wp-content/uploads/2015/01/Renner_Grenze-und-Ereignis.pdf)
- Ryan, Marie-Laure (2014): “Space.” In: *The Living Handbook of Narratology*, hrsg. von Peter Hühn, Jan Christoph Meister, John Pier und Wolf Schmid. Hamburg: Universität Hamburg.  
<https://web.archive.org/web/20221210041153/https://www-archiv.fdm.uni-hamburg.de/lhn/node/55.html>

## Literaturhinweise

- Schumacher, Mareike (2023): *Orte und Räume im Roman. Ein Beitrag zur digitalen Literaturwissenschaft*. Berlin, Heidelberg: Springer.  
<https://doi.org/10.1007/978-3-662-66035-5>.
- Schumacher, Mareike (2022): “‘Wie Wölkchen Im Morgenlicht’ – Zur automatisierten Metaphern-Erkennung und der Datenbank literarischer Raummetaphern LaRa.” In: *DHd2022: Kulturen des digitalen Gedächtnisses. Konferenzabstracts*, S. 232–236.  
<https://doi.org/10.5281/zenodo.6304590>.
- Viehhauser, Gabriel (2020): “Zur Erkennung von Raum in narrativen Texten. Spatial Frames und Raumsemantik als Modelle für eine digitale Narratologie des Raums.” In: *Reflektierte Algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt*, hrsg. von Nils Reiter, Axel Pichler, and Jonas Kuhn, S. 373–388. Berlin und Boston: De Gruyter. <https://doi.org/10.1515/9783110693973-015>.
- Viehhauser, Gabriel und Florian Barth (2017): “Towards a Digital Narratology of Space.” In: *Digital Humanities 2017. Conference Abstracts*, S. 643–646. Montréal.  
<https://web.archive.org/web/20230615201230/https://dh2017.adho.org/abstracts/DH2017-abstracts.pdf>