

Shared Tasks in DH

Systematic Analysis of Narrative Texts through Annotation (SANTA)

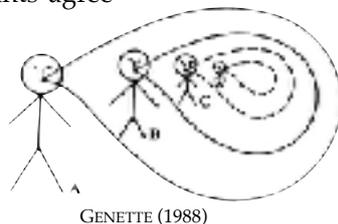
Our proposed DH shared task works in two phases:

Phase 1: SANTA (starting summer 2017)

Participants develop annotation guidelines for narrative levels and the narrator position to them (homo-/heterodiegetic) independently, given a diverse corpus of narrative texts. In spring 2018, participants will submit their annotation guidelines to the organisers, together with annotations on a provided selection of texts. Participants are then asked to annotate the same texts according to annotation guidelines of another participant. In addition, students coordinated by the organisers will annotate using the guidelines.

Participants will be invited to a **workshop**, to discuss the different guidelines. Due to the re-distribution of the guidelines and the student assistants, the discussion can be informed by an empirical evaluation of inter-annotator agreement. The result of this workshop will be a set of guidelines that the participants agree upon.

Result: Clearer, more explicit and unambiguous definition of narratological phenomenon; evaluated annotation guidelines.



Annotation Phase

These resulting guidelines will then enter the annotation phase, in which student assistants will annotate a larger corpus. **Result:** An annotated corpus

Phase 2

The goal in Phase 2 is to foster the development of systems to automatically detect narrative levels in texts. A large part of the annotated corpus will be released early, as training and development material. A small part will be withheld, in order to be used as (unseen) test data later on. Since test data will only be available a short time before submission deadline, tuning systems to a specific text/corpus becomes unfeasible.

The submitted automatic annotations will be compared quantitatively, using standard evaluation metrics. A **workshop** (likely in coordination with LaTeCH) will take place in which all participants describe their systems.

Result: Experiments on the best approaches to automatically detect narrative levels; meaningful comparison of different approaches.

Frequently Asked Questions

What is a shared task?

Shared tasks are a research framework popular in natural language processing (NLP). In NLP, different participant teams develop systems that all solve the same task. Due to previously set standards and evaluation data sets, the performance results are directly comparable. This gives insight which systems are able to solve a task.

Why should I take part in this?

Because working collaboratively on annotation guidelines fosters the advancement of definitions of narratological concepts. This in turn is prerequisite for manual and automatic large-scale analyses. In addition, submitted and created work will be published and properly citable.

Can I participate with a group?

Yes, absolutely. You can even integrate this with a seminar and participate with a student group.

Why narrative levels, focalisation is more interesting?

Three major reasons: First, narrative levels are everywhere in a text and also **frequently decidable** (while focalisation, for instance, is only rarely marked). Second, it is **ubiquitous**: Most narrative texts and even many non-fiction texts (e.g., longer pieces in magazines) contain multiple narrative levels. Third: We believe that switches between narrative levels can be **detected automatically**.

Can I participate in only one phase?

Yes. Phase 1 requires no technical skills at all. Conversely, participation in Phase 2 does not require knowledge of narratology.

Who we are

Dr. **Evelyn Gius**, German language and literature, Hamburg University

Dr. **Nils Reiter**, computational linguistics/digital humanities, Stuttgart University

Dr. **Jannik Strötgen**, computer science, Max Planck Institute for Informatics in Saarbrücken

Dr. **Marcus Willand**, German language and literature, Stuttgart University

Advisory Board

M.A. **Janina Jacke**, Hamburg University; Prof. Dr. **Fotis Jannidis**, Würzburg University; Prof. Dr. **Jonas Kuhn**, Stuttgart University; Prof. Dr. **Jan Christoph Meister**, Hamburg University



<https://sharedtasksinthedh.github.io>