

the insight that could be extracted by their carefully crafted school diaries or soldier diaries written as part of a veteran support programme, reporting reading years after they occurred.

**Case 2 – Heterogeneous comparative studies.** Having a shared definition of reading enables comparative studies among heterogeneous sources. For instance, an ongoing comparative study on popular reading brings together a) long, matured reviews of books from online reading groups, with b) short impulsive comments on webcomics "issues". The design of heterogeneous comparative studies also requires careful considerations (Benatti et al, 2021) about, e.g., the practices of the audiences, maturity of the experience or type of information provided, connected by the form of prompting.

To summarise, data interoperability is a precondition, but it is not sufficient: a common language of the phenomenon does not exhaust the information needed for interpretation. Specifically, in the vision of step further toward research interoperability - as the synergy between studies in terms of agenda and findings contributing to the knowledge of a common phenomenon – a computer-readable epistemology of research case studies should be included as a part of their output.

In both cases, the data do not reflect these considerations, emerged only by talking directly with the involved researchers. To address this issue, we designed two un-planned ontologies: Experience & Observation (E&O) ([https://github.com/modellingDH/odp\\_experience](https://github.com/modellingDH/odp_experience)) and Profiles, Groups & Communities (PGC) (<https://github.com/modellingDH/profile-group-community-odp>).

E&O is currently in the READ-IT contribution platform. The first application of E&O is to document the different modalities we use in crowdsourcing of experience of reading, e.g. through postcards, webforms and chatbot conversations (<https://readit-project.eu/contact/contribute-to-read-it/>). Secondly, E&O describes the relations between activities and prompting of reading experiences, typical of the different sources (e.g., topics of the questions asked, time from reading). This use of E&O helped identify nine recurring patterns exposing new, unexpected similarities and differences between case studies that can be evaluated through objective metrics (Antonini et al, under review).

PGC complements the characterisation of readers introduced by REO, with the difference between reader and status of the reader at the time of reading, by introducing the missing social dimension. Specifically, PGC addresses the specific reader profiles in terms of, e.g., linguistic competencies or core values, their belonging to social bodies (religious groups, political parties) or community of practice (as reading groups or professions such as editors or scholars).

This contribution addresses the epistemic gap in how data about experiential research, which is a barrier in

moving beyond data interoperability toward a conceptual integration between different research case studies. This gap is addressed by introducing two models used to explicit essential information about research case studies and sources to complement and give context to research data. Both gap and solution are discussed based on actual case studies on reading.

## Bibliography

- Our-Vial, B. and Antonini, A.** (2021) Reconnecting with the Evolving Journey of Reading. *Cultural Practice*, Re-Connect, 2021, Last access September 2021 <https://www.culturalpractice.org/article/reconnecting-with-the-evolving-journey-of-reading>.
- Vignale, F., Benatti, F. and Antonini, A.** (2019). Reading in Europe - Challenge and Case Studies of READ-IT Project. In: *Digital Humanities Conference 2019*, 9-12 Jul 2019, Utrecht, Netherlands.
- Vignale, F., Antonini, A. and Gravier, G.** (2020). The Reading Experience Ontology (REO): Reusing and Extending CIDOC CRM. In *proceedings of DH2020*. 2020, Ottawa, Ca.
- Antonini, A., Suárez-Figueroa, M. C., Adamou, A., Benatti, F., Vignale, F., Gravier, G. and Lupi, L.** (2021). Understanding the phenomenology of reading through modelling. *Semantic Web*, 12(2), 191-217.
- Benatti, F., Norrick-Rühl, C. and Antonini, A.** (2021) Reading Popular Culture Offline and Online: Outlining a Comparative Study of Reading Experiences Between Webcomics and Twenty-First Century Book Club Choices. *EADH2021*, September 2021.
- Antonini, A., Adamou, A., Suárez-Figueroa, M. C. and Benatti, F.** (under review) Experiential Observations: an Ontology Pattern-based Study on Capturing the Potential Content within Evidences of Experiences. Special issue on Cultural Heritage 2021, *Semantic Web – Interoperability, Usability, Applicability* an IOS Press Journal (under review), URL: <http://semantic-web-journal.net/content/experiential-observations-ontology-pattern-based-study-capturing-potential-content-within>.

## Salience in Literary Texts: A Combined Approach to the Relevance of Passages

**Arnold, Frederik**

frederik.arnold@hu-berlin.de  
Humboldt-Universität zu Berlin, Germany

## Fiechter, Benjamin

fiechtbe@hu-berlin.de

Humboldt-Universität zu Berlin, Germany

## Gius, Evelyn

evelyn.gius@tu-darmstadt.de

Technical University of Darmstadt, Germany

## Jäschke, Robert

robert.jaeschke@hu-berlin.de

Humboldt-Universität zu Berlin, Germany

## Martus, Steffen

steffen.martus@rz.hu-berlin.de

Humboldt-Universität zu Berlin, Germany

## Vauth, Michael

michael.vauth@tu-darmstadt.de

Technical University of Darmstadt, Germany

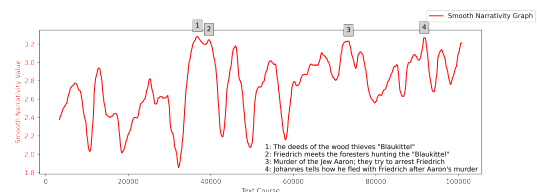
## A combined approach to the relevance of text passages

In this contribution, we want to outline insights that arise from combining two distinct approaches to literary texts that analyse the relevance of specific text passages. We have been working on the identification of the **narrativity** represented in literary texts as well as on the **quotation** of the texts in research to identify passages especially relevant from a hermeneutical perspective. This now allows us to explore whether the structures and patterns that are emerging from these two approaches can be related to each other in a meaningful way.

### Narrativity as textual relevance criterion

Our identification of narrativity of literary prose texts is based on the annotation of events. By considering the core features of events in narrative theory (i.e., being a state, a process in time and change of state) we classify each verbal phrase in a text as change of state, process event, stative event or non-event.<sup>1, 2</sup> To enable measuring narrativity, this categorical scaling is transposed into a numerical scaling reflecting the degree of narrativity of the event categories. In accordance with a narrative theory understanding of events, we determine the narrativity of the annotation categories with the values 7 (change of state),

5 (process event), 2 (stative event), and 0 (non-event). By additionally smoothing the narrativity value we are able to model the narrativity of a text as a graph.<sup>3</sup> Figure 1 shows the narrativity graph for the novella *Die Judenbuche* by Annette von Droste-Hülshoff which serves as an example for our approach.

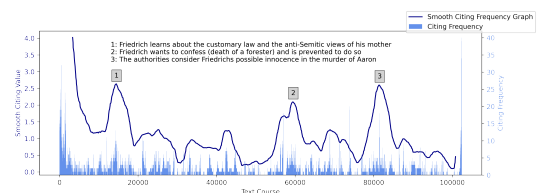


**Figure 1.**  
*Narrativity score*

### Key passages: quotation as textual relevance criterion

We consider key passages as parts of a literary text that are especially relevant for interpretation and can differ in length from only a few words to one or more paragraphs. To learn which parts of a text are more relevant than others, we rely on expert knowledge, which we obtain from numerous interpretations of a literary work containing quoted passages. This is a new approach in text and literature studies, that has not been theoretically founded yet; though the term “Schlüsselstelle” (key passage) and equivalents are used regularly in text interpretations in German language.<sup>4</sup>

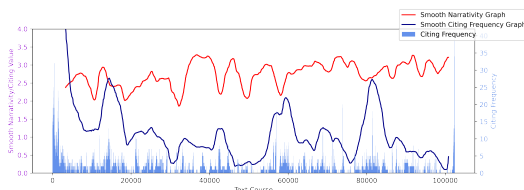
For this study, we limit ourselves to a quantitative view. We have analysed 44 interpretations of *Die Judenbuche*, all in German language, published between 1995 and 2015 and identified quoted passages with a Python tool for quotation detection in fictional texts.<sup>5</sup> Figure 2 visualises the identified quotations over the course of the text; the histogram shows quotation frequency and the graph the smoothed frequency for each verbal phrase identified during the event annotation. Notably, the beginning and the end are quoted most frequently, together with three other frequently quoted passages.



**Figure 2.**  
*Quotation frequency*

## Combining the approaches: frequently quoted (key) passages and narrativity

By combining exploration of narrativity and quotation frequency (cf. Figure 3) we can explore whether a passage is referred to as one that is mainly interesting for the storyline or for the interpretation. Passages with a high narrativity score are particularly important for the plot and the comprehension of the text, while passages with a low narrativity value more often contain dialogue or narrator comments in which interpretation proposals are already made that are taken up in literary studies texts. For passages with a medium narrativity value, potential interdependencies are difficult to determine on the basis of only one text, but we are aiming to obtain more detailed knowledge on this in the future, including also non-frequency based analyses of references. Also, the beginning and the end of the text seem to be quoted in a different manner. Here quotation frequency and narrativity seem to be connected only loosely. Instead, these borders of the text seem to be used mostly to provide a framework for interpretations, in which the interpreters select the most interesting passages for their intent.



**Figure 3.**  
*Narrativity score and quotation frequency*

## Outlook

While these findings point out how the quotation of text passages may relate to their narrativity, they should be evaluated against a broader corpus of texts. There, the classification of functions of quotations is the most interesting aspect. We assume that plot-orientated quotations in the secondary literature correlate with higher narrativity, whereas passages quoted in order to develop a more comprehensive interpretation of the text display less narrativity. For evaluating this, we plan to combine our automated analysis of narrativity with the automated detection of key passages.

## Bibliography

- Arnold, F. and Fiechter, B.** (2022). Lesen, was wirklich wichtig ist: Die Identifikation von Schlüsselstellen durch ein neues Instrument zur Zitatanalyse. *DHd2022: Konferenzabstracts*. <https://doi.org/10.5281/zenodo.6327917> (accessed 11 April 2022).
- Arnold, F. and Jäschke, R.** (accepted). Lotte and Annette: A Framework for Finding and Exploring Key Passages in Literary Works. *Proceedings of the Workshop on Natural Language Processing for Digital Humanities at ICON 2021*.
- Gius, E. and Vauth, M.** (2022). Inter Annotator Agreement und Intersubjektivität. *DHd2022: Konferenzabstracts*. <https://doi.org/10.5281/zenodo.6328209> (accessed 11 April 2022).
- Vauth, M. and Gius, E.** (2021). Richtlinien für die Annotation narratologischer Ereigniskonzepte. *Zenodo*. <https://doi.org/10.5281/zenodo.5078174> (accessed 11 April 2022).
- Vauth, M., Hatzel, H. O., Gius, E. and Biemann, C.** (2021). Automated Event Annotation in Literary Texts. *CHR 2021: Computational Humanities Research Conference*. Amsterdam, pp. 333–45. [http://ceur-ws.org/Vol-2989/short\\_paper18.pdf](http://ceur-ws.org/Vol-2989/short_paper18.pdf) (accessed 11 April 2022).

## Notes

1. For a detailed explanation of the manual annotation on which the data used in this contribution is based cf. Vauth and Gius, 2021.
2. For a description of classification of events as well as the automation approach and results cf. Vauth et al., 2021.
3. For a discussion of the adequacy of this implementation for literary studies, especially with regard to intersubjectivity, cf. Gius and Vauth, 2022.
4. For more details on key passages and the aim of the project cf. Arnold and Fiechter, 2022.
5. For a detailed explanation, cf. Arnold and Jäschke (accepted). Source code available at: <https://scm.cms.hu-berlin.de/schlueselstellen/lotte>.