An Interactive Visualization for the Analysis of Annotated Text Variance in the Legendary *Der Heiligen Leben, Redaktion*

Martin Baumann*, Markus John*, Hermann Pflüger*, Cornelia Herberichs†, Gabriel Viehhauser‡, Wolfgang Knopki§, and Thomas Ertl*

*Institute for Visualization and Interactive Systems, University of Stuttgart, Stuttgart, Germany
E-mail: firstname.lastname@vis.uni-stuttgart.de

†German Department, University of Fribourg, Fribourg, Switzerland
E-mail: cornelia.herberichs@unifr.ch

‡Institute of Literary Studies, University of Stuttgart, Stuttgart, Germany
E-mail: viehhauser@ilw.uni-stuttgart.de

§Information Center, Hochschule für Technik, Stuttgart, Germany
E-mail: wolfgang.knopki@hft-stuttgart.de

Abstract—The medieval legendary *Der Heiligen Leben, Redaktion* exists in two variants. As part of an ongoing digital edition project, the difference phenomena between these variants are manually annotated and categorized. We present an interactive, comparative visualization approach for the exploration of these annotation data with the goal of analyzing the nature of text variance in this work. The approach combines a set of coordinated close and distant reading views in an overview+detail design and provides means for category-filtering and detailed inspection of the data. We conducted a task and requirement analysis in cooperation of the involved medievalists and visualization researchers and sketch two scenarios of usage.

Index Terms—Visual Text Comparison, Text Annotation Visualization, Digital Humanities

I. INTRODUCTION

Der Heiligen Leben is a medieval German legendary, that is to say a collection of saints' legends. It emerged around the end of the 14th century, and in the 15th century, it became the most popular and widespread legendary, both as manuscripts and later as early prints until the Early Modern Age. In its original form, it was comprised of about 250 legends. According to Kunze, the precise number of the original extent is unknown [1]. Then, at the beginning of the 15th century, this collection was enriched with about 150 legends and combined with a martyrologium, a list of martyrs, saints and beati ordered by their anniversaries. This extended, combined work came to be known as *Der Heiligen Leben, Redaktion*. Only a few decades later, a new edition of this work was created, where many of the legends were retold with significant changes, like the addition of new details, stylistic changes, shortenings or rearrangement of text passages. Henceforth, we will use the abbreviation HLR when we speak of the pair of these two variants of *Der Heiligen Leben, Redaktion*. There

is a two-volume edition of *Der Heiligen Leben* by Brand et al. [2], [3], and for a scholarly introduction consult the work of Williams-Krapp [4, p. 315-338].

Presumably, the reasons for the differences between the two versions of the HLR can be separated into two groups – at the moment, however, this is just a hypothesis, one that, among others, the edition project described below seeks to confirm or reject on a broad data basis. First, there might be extra-literary reasons: Since these were also functional texts, different contexts of usage implied different requirements to the text form. For example, as Williams-Krapp remarks, the prayers at the end of the legends were significantly lengthened, possibly due to the usage in a liturgical context [4, p. 336]. Second, there might be literary reasons resulting mainly from a series of tensions. For example, on the one hand, the legends should encourage the common believer to an imitation of the saint's exemplary life, while on the other hand, this saint should be presented as someone who is extraordinary and performs extraordinary deeds. On this subject, see Köbele [5]. Or as another example: While a legend should speak for itself (and the humble author should step behind the object of his text), the effective presentation of such an extraordinary and exemplary life required the usage of quite some rhetoric machinery. Hence, these tensions and different contexts of usage could be the reason for the repeated retelling of the texts.

The systematic analysis of the edition tendencies between the two versions of the HLR appears as an important desideratum in the field of medieval studies. On the one hand, the literary genre of legends has advanced to a field of intensive research in medieval literary studies over the course of the last ten years (see for example the monograph by Weitbrecht et al. [6] and further works by these authors and by S. Köbele). On the other hand, however, the *legendary* as a collective liter-

Fig. 1: The approach's main window with its components: category bar A, text overview B, phenomenon bar C, text detail views D1 and D2, and drop-down menu E. All categories are selected, phenomenon no. 170 is activated.

ary form in its own right, has by now hardly been researched. This hold true for the HLR in particular, which is the most widespread legendary of the middle ages. The two medievalists among the authors have started a digital edition of the HLR that aims at contributing to the task of researching its edition tendencies. One basic idea of this project is to systematically *categorize* the occurrences of differences between the texts – such an occurrence we henceforth will call a *phenomenon* – and to manually *annotate* a phenomenon's respective sets of words – or *tokens* – in both texts, that is to tag these sets with the metadatum of the phenomenon category that their common phenomenon belongs to. These phenomenon annotations for all of the legends, together with further metadata such as named entities, could then be used for further inquires by the medievalist community.

As a simple example, consider the legend of Ambrosius, bishop of Milan in the 4th century and patron saint of beekeepers. There is an episode in this legend telling how Ambrosius forced the Roman emperor to make atonement for a massacre he had committed on the people of Thessaloniki. In the second version of the HLR, the emperor is never called by his name. In the first version however, the author speaks of “Falentinus”, the Roman emperor Valentinian. This is incorrect, since this massacre was committed by Valentinian's successor, Theodosius II. Probably, the author of the first version inserted the name of Valentinian simply because he already appears in an earlier passage of the legend. Now, it would be interesting to analyze this phenomenon over larger portions of the text variants and in conjunction with other insertions of person names and also with co-occurring phenomena of other categories.

One challenge of the endeavor of a digital edition with rich

annotations results from the amount and complexity of the data that the medievalists work with and produce in the course of their annotation task. The HLR versions are, together, comprised of roughly 2 585 000 tokens (including stop words), and the authors estimate that by the end of the annotation phase of the edition project, about 120 000 phenomenon-occurrences (and hence twice as many annotations) will have been marked by the annotators. As for the complexity, each phenomenon consists of two related token sequences, a series of tags marking categories, that can be linguistically complex in themselves, and further optional components; details will be given in section III. In order to cope with these data and exploit their full potential for the analysis of the HLR, the medievalists and the visualization researchers comprising the author list of this paper cooperated to develop an approach for the visual exploration of the annotation data. It supports the direct visual comparison of each legend's two variants by displaying the texts and the annotations on different levels of proximity (close and distant views) and by providing different means of navigation throughout. The details will be given in section IV. The derivation of the approach's requirements (see section III) as well as its assessment (see section VI) were conducted in close collaboration between the project partners.

To sum up, the contribution of this paper is the introduction of an approach that allows to visually and interactively explore the differences of the two variants of the HLR that are encoded in a set of manually derived phenomenon annotations. The metadata that these annotations carry consist of several components, most notably a series of tags that characterize the categories of the respective phenomenon semantically. Following an overview+detail design in the sense of Cock-

burn [7]), the approach combines a series of coordinated views (see Figure 1): detail views that allow for a close reading of the text variants and overviews that allow for a distant reading of the annotation data in the sense of Moretti [8].

We will show in section II that on the one hand, similar problems on similar data sets have been tackled within the visualization community, but that on the other hand, none of these approaches can provide the functionality needed to deal with the data and tasks of our project. We will sketch a few possible ways of usage in section V, and report on the feedback we got on several occasions from the medieval studies and digital humanities communities in section VI. Finally, we will sketch some promising paths of future development of the approach in- and outside of the frame of the edition project in section VII.

II. RELATED WORK

When compared to similar projects, the characteristics of our data are that first, we have only two text variants, that second, the differences between these two variants are not massive in terms of number and length, but that third, the differences are encoded in set of rich, manually generated annotations that, among other aspects, categorize the respective difference phenomenon semantically. Based on these characteristics, we work out in the following the commonalities and differences between our approach and the current state of the art in the visualization of text differences. In doing so – especially in the last, more elaborate example – we refer to several features of our approach with a terminology that is only explicated in section IV.

The *Versioning Machine* of Schreibman allows to compare any number of TEI-encoded texts variants [9]. Similarly to our approach, metadata can be added manually to the texts. However, the visualization support consists only in highlighting corresponding text passages.

More visualization support is provided by the *Juxta Commons* tool [10]. It allows for the automatic collation of multiple text variants and can display two of them side by side together with a middle column of bars connecting text areas that contain difference annotations – similarly to what the text overview arcs of our approach in conjunction with its phenomenon bar provide. However, there is no categorization of the occurrences of text differences available, the middle column visualization gives only a local view and may quickly become quite fussy with many overlapping occurrences, and the only abstracting overview visualization is provided by a histogram.

The *VarifocalReader* by Koch et al. [11] is another example of an approach that allows to display two annotated text variants side by side. It provides features to easily navigate within long texts and shows different aggregations of the texts at several levels of hierarchy. However, the main aspect of our annotation data, namely that two annotations are always combined to form a common phenomenon of difference, is not supported.

There are a number of approaches, that have one of their main strengths in the scalability with respect to the number

of variants to compare – an aspect that is not of importance in our case. With their *ShakerVis* approach, Geng et al. [12] compare 32 retranslations of a scene of Shakespeare’s *Othello* in order to identify those segments, that provoked particularly strongly differing translations. They use dimension reduction techniques on vector representations of the segment translations, compute similarity measures between them and display the results in an array of coordinated heat maps, parallel coordinate plots and scatter plots.

Asokarajan et al. [13], [14] devised a pixel-based focus+context approach to compare a large number of text witnesses against a base version per word, line and page. They support the text scholars’ task to explore patterns of variation across a base text’s lemmata in comparison with a given witness, and across a range of witnesses for a given lemma of the base text. They do this by displaying in the focus area a matrix of lemmata versus witnesses, where each cell encodes in its color the edit distance between a specific lemma of the base variant (column) and the corresponding lemma of a specific witness (row). In the context area, they accumulate these edit distances for each witness across lines (close context) or pages (distant context) of the base variant. Their approach scales well to a large number of text witnesses, but they offer no support for a categorization of the differences.

Jänicke et al. [15] suggest a visualization approach for the distant reading of a corpus of 24 translations of the bible. The corpus lends itself particularly well here, because each translation is structured on four levels of hierarchy: work, “books”, chapters, and verses. They assess the (dis-)similarity of a selection of translations with respect to a deliberate majority level of the whole corpus as the number of differing words. The result of this assessment on one of the hierarchy levels mentioned above is then displayed in a series of fingerprint visualizations. Again, while this approach scales well with the number of variants, the specific kinds of differences between the translations cannot be accessed, and as a decidedly distant view, it does not offer access to the texts themselves.

Going to the other end of the zooming range, there is a class of approaches dealing with text variation at word level, such as Jänicke’s variant graph approach *TRAViz* [16] – a progression from the older approaches of *CollateX* by Dekker and Middell [17] and *Stemmaweb* by Andrews [18] – the *Word Tree* by Wattenberg and Viegas [19], the *WORDGRAPH* by Riehmann et al. [20], or the approach by Silvia et al. [21] that adapts the storyline metaphor to variant graphs.

Some further examples related to text comparison and re-use encompass the approach for the assessment of plagiarism by Riehman et al. [22], the text similarity detection approach of Abdul-Rahman et al. [23], the set of close and distant reading visualizations suggested by Jänicke et al. [24], or the *News Auditor* approach by Behrisch et al. [25] that allows to explore news articles forming clusters around specific “stories”.

Finally however, the works closest to the our approach in terms of data, problem and proposed visualization solution are those of Jänicke and Wrisley [26], [27]. Here, they propose a visual analytics approach to semi-automatically align highly

differing variants of medieval lyric texts. Their approach encompasses representations at different levels of hierarchy (whole text, lines, words), and in particular a view that they characterize as “meso” in that it allows to read the texts and at the same time and in the same view displays information involving distant parts of the differing variants. Here, they employ stream graphs to connect aligned lines of text across different variants. This bears a strong resemblance to the combination of text detail views and text overviews in our approach. Furthermore, they include previews of the word level alignments as word-sized graphics next to the text lines, which seems conceptually similar to our line glyphs in the text overview. They incorporate variant graph visualizations as pop-ups, similarly to our phenomenon detail window, and finally, their macro perspective – a series of lines connecting abstractions of aligned verses – corresponds to the text overview of our approach.

There are, however, also significant differences to our approach. First, with respect to the data and tasks, Jänicke and Wrisley support the high-level task of analyzing the “mouvance”, that is to say patterns of text re-use, in highly instable medieval texts, possibly between many variants. This results in the low-level task of semi-automatically aligning related text fragments across variants. Hence, while their subject of analysis are the configurations of similar passages across many, highly instable variants, ours are the exact nature and distribution of the differences between two relatively similar variants, and these differences are, in our case, not to be detected automatically but given as rich, manually created annotations.

Second, with respect to the visualization solution, they separate the macro and meso perspective, whereas in our approach, the text overview is integrated between the text detail views, more in the spirit of the Juxta Commons approach mentioned above. While our version may not scale as well to many text variants or very long texts – as Jänicke and Wrisley have to deal with – it has the advantage that for phenomena, whose annotations are to be found in very distant parts of the texts, the connection of these parts can still be fully shown within the visual field of the text views. Then, while the phenomenon detail window in our approach allows to inspect the data more closely, it does not – as Jänicke and Wrisley do with their variant graphs – provide a conflating view on distant, related lines of text. This is a functionality that we rather offer by means of the reduced viewing mode in conjunction with the filtering options in our approach. Finally, the phenomenon bar of our approach, which gathers the phenomena as central units of study in their own view, independent of the texts, finds no equivalent in the approach of Jänicke and Wrisley – not least since the data and tasks of the approaches are only partly comparable.

III. DATA, TASKS AND REQUIREMENTS

The contributing medievalists set up a series of annotation guidelines and manually annotated some of the HLR’s legends in order to provide the data for the contributing visualization

TABLE I: The phenomenon categories with color scheme.

0	additional adjectives	6	variant name
1	additions within a sentence	7	additional person
2	additional episode	8	additional place name
3	synonyms	9	syntactic rearrangement
4	similar semantics	10	genera verbi (passive/active)
5	changed semantics	11	oratio (direct/indirect speech)

researchers to build their first prototypes. The annotation data is structured as follows (an example can be seen in the phenomenon detail window in Figure 2):

Fig. 2: Pop-up window with details of phen. no. 13c: categories, certainty indicators, annotation segments, comments.

- A difference phenomenon is comprised of two annotations, one in each text variant, and the relation between them.
- Each phenomenon is characterized by an identification number.
- The metadata a phenomenon carries consist of one or more categories as tags, a number indicating how certain the annotators are in their allocation of the respective category, the two token sets of the phenomenon’s annotations, and optional comments to each of the annotations.
- The text segments of different phenomenon annotations are allowed to overlap.

Right now, we have twelve categories of phenomena, they are displayed in Table I. We should mention however, that the number and content of these categories are only provisional and may very well change and evolve as the edition project matures further.

The tasks that the contributing medievalists as well as typical prospective analysts of the data set want to pursue encompass:

- T1 Inspect the specific content of phenomena in the textual contexts of their annotation segments.
- T2 Analyze the distribution and textual length of phenomena of a specific category over the course of a legend.
- T3 Analyze patterns of co-occurrences of phenomena of different categories.

In order to support these tasks, our approach should fulfill the following design requirements:

- R1 Close reading views: Show both texts simultaneously together with the annotation data (token sets and categories) and indicate annotations with overlapping token segments and/or multiple categories. Show further details of the annotation data on demand.
- R2 Distant reading views: Give an overview over the annotation data for an entire legend that shows their distribution over the texts and allows for direct interaction. The close and distant reading views should be interlinked for browsing the texts and the annotation data.
- R3 Filtering: Allow for the selection of categories, such that only those annotations tagged with these categories and only those portions of text containing segments with such annotations are shown.

A comprehensive survey over the problem of supporting close and distant reading by means of interactive visualizations can be found in Jänicke et al. [28], [29].

IV. APPROACH DESCRIPTION

In order to facilitate the exchange with our project partners, we implemented the approach as a web-based application using the D3 library [30]. The word files that were annotated by the medievalists were converted to an XML-file via a python-script, which was then fed to the tool.

The overview part of our overview+detail design consists of three elements (see Figure 1), the *category bar* A, the *text overview* B, and the *phenomenon bar* C. These elements allow for a distant reading of the annotated phenomena in the two text variants (R2). The detail part, which allows for a close reading of the texts (R1), consists of the two *text detail views* D1 and D2. Further details about the annotations are provided in the *phenomenon detail window*, popping up upon demand, that can be seen in Figure 2 (R1). A drop-down menu provides a category legend and some adjustment levers (see Figure 1 E).

A. Overviews

In the category bar, the analyst can choose which annotations and phenomena should be highlighted in the respective color of their category in the other views, and hence this filtering action determines the appearance of all other views (R3) (see Figure 1 A). The length of one of the colored sub-bars linearly encodes the number of phenomena in the respective category and therefore, the category bar indicates the distribution of phenomena over the different categories. The description of these categories can be seen in a tooltip or in the legend given in the drop-down menu (see Figure 1 E). In this menu, besides looking at the category legend, the analyst may toggle the *reduced viewing mode* of the text detail views (see subsection IV-B), open up some data statistics (which we shall not further discuss here) and choose the font size of the texts.

The text overview in the middle of the main window (see Figure 1 B) consists of two columns of line glyphs and – between them – a collection of arcs that connect some of these glyphs. Each line glyph (see Figure 3) corresponds to

Fig. 3: Blowup of the line glyphs. See Figure 1 B for context.

a line of text; its length encodes the logarithm of the number of tokens in the respective line, and the colored subparts indicate which of these tokens carry an annotation. We chose to use a logarithmic scale here because the line lengths vary substantially, and a relatively uniform width of the line glyph columns appeared as more desirable than the benefit of being able to directly compare line lengths on a linear scale.

The two line glyph columns each represent one of the two text variants in a condensed form, such that the analyst can grasp the distribution of phenomena within the texts at once. Two line glyphs are connected by an arc if their respective lines of text contain annotations that belong to the same phenomenon. A gray selection rectangle over each line glyph column indicates those lines, whose text can be seen in the text detail views.

The phenomenon bar at the bottom shows each phenomenon as a rectangle colored according to the respective categories, together with its identification number (see Figure 1 C and Figure 4). If the respective categories are not chosen in the cate-

Fig. 4: Cut-out of the phenomenon bar. Phen. 13c is activated.

gory bar, the rectangles are grayed out. This view is in a sense orthogonal to the text overview in that it emphasizes rather the phenomena than the text by attributing each phenomenon the same visual weight, independent of the extensions of its annotations in the two texts. Two small red bars at the top and at the bottom of the view indicate those phenomena whose annotations can be seen in the left and right text windows right now. The entire view can be hidden upon request.

In an early prototype, we experimented with a different layout of the phenomenon bar (see Figure 5). Instead of having one row of phenomenon rectangles that can be scrolled horizontally, we implemented a matrix of rectangles that was to be read row-wise. The advantages of the latter variant would be that it requires less scrolling, that the scrolling would

Fig. 5: Alternative, matrix-like design of the phenomenon bar.

be vertical rather than horizontal, and that more phenomena (maybe even all, in the case of only a few text differences) could be displayed simultaneously – which would provide a more comprehensive overview over the phenomena. However, the disadvantages would be that first, if you would want to display a lot of phenomena, you would have to dedicate a large amount of screen space to the phenomenon view, and second, that spatial proximity of phenomena (say, in the same column) might wrongly suggest spatial proximity in the text. Therefore, we went with the alternative, row-wise layout for now, but other layouts or focus+context techniques might be tried and be given as an option to the analyst.

B. Detail Views

In the text detail views, the analyst can read portions of the two text variants side by side and can see the annotations belonging to a certain phenomenon in their respective context (see Figure 1 D1 and D2, and Figure 6). The annotations

Fig. 6: The text detail window for one text variant: annotation underlines, one activated phenomenon (gray highlight), and text segments carrying overlapping annotations (red frames).

whose categories were selected in the category bar are shown by a colored underline. If the text segments of different annotations overlap, a red frame around the overlapping region invites the analyst to inspect the passage more closely by bringing up the phenomenon detail window. Upon activating an annotation by mouse hover or click (see subsection IV-D), the respective text segment is indicated by a gray highlight.

In an early stage of development, two other design alternatives to the colored underlines in the text detail views were considered. In the first, we put highlights on the text that – in the case of multiple categories – were split horizontally (see Figure 7a). This way, it would be possible to encode uncertainty information in the height of the highlight segment of the respective category (see Figure 7b). While being a good solution for long annotations with many categories, this version had the disadvantage that the color borders of an annotation with multiple categories could be easily confounded with the annotation segment borders of multiple annotations with one category.

Fig. 7: Design alternative: highlights as annotation marks.

In the second alternative, we put highlights that were split vertically (see Figure 7c). This avoided the problem mentioned above, but one could hardly show more than three simultaneous categories (see Figure 7d), and would not have the option to encode uncertainty information for multiple categories. However, both highlighting versions would render the text very colorful and busy, and therefore, we went with the underline design instead in order to keep the texts more readable.

In the simplified text detail view, only those lines of text are shown, that contain annotations of the categories chosen in the category bar (R3, see Figure 8). Lines of text over which an

Fig. 8: The simplified text detail views. Category 3 (synonyms) is selected. Activated phenomenon: “haist” vs. “genant ist”.

annotations spans are grouped and marked by a gray rectangle (which changes to red if the annotation is activated).

A right click on an annotation in one of the text detail views or on a phenomenon in the phenomenon bar opens the phenomenon detail window (see Figure 2). Here, analysts can first select which phenomenon to show – if several of them

overlap on the respective text segment. Then, they are shown the list of categories the phenomenon belongs to, the token sets of the annotations in the two text variants, and – if provided – comments and certainty indicators.

C. Color Encoding

The color coding of categories should be done such that the texts would be still well legible, that the different categories could be easily discriminated, and that there is a correspondence between color proximity and conceptual proximity of the categories. The colors were chosen by employing the palette of *colorbrewer* [31], and choosing a set of reddish colors for supplements (categories 0 to 2, 7 and 8), blueish colors for semantic categories (3 to 5), and colors in a maximal distance for categories 6, 9, 10 and 11 (see Table I). For phenomena belonging to multiple categories, underlines and phenomenon rectangles are colored by a gradient between the first two category colors on the list, hence inviting the analyst to inspect the phenomenon more closely in the phenomenon detail window.

D. Interactions

The different parts of the approach react differently to the three mouse actions of hover, click and right click – these interactions are summarized in Table II. A right click on a

TABLE II: Interaction effects: tool elements vs. mouse actions.

	hover	click	rightclick
phen. (bar)		activate & focus phen.	show details
annot. (text)	activate phen.	activate & focus phen.	show details
line glyphs		move text view	

rectangle in the phenomenon bar or on an annotation in one of the text detail views opens the phenomenon detail window. Hovering over an annotation in a text detail view *activates* the corresponding phenomenon, which has the following consequences:

- Both annotations belonging to it are highlighted in gray in the text detail views.
- The phenomenon’s identification number in the phenomenon bar is highlighted in gray.
- In the text overview, the arc connecting the respective line glyphs is highlighted in red.

Clicking on an annotation in a text detail view or on a rectangle of the phenomenon bar activates and *focuses* this phenomenon. This means that both text detail views are scrolled automatically such that the annotations belonging to the phenomenon appear on top (and the selection rectangles in the text overview are also adjusted accordingly), and the phenomenon bar is scrolled such that the focused phenomenon appears to the left. Clicking somewhere on one of the text overview’s line glyph columns moves the respective selection rectangle and the respective text detail view is scrolled to this selection.

V. USAGE SCENARIOS

In order to demonstrate some simple paths of usage for our approach and to illustrate some of the tasks mentioned in section III, we will now sketch two examples of usage scenarios.

In the first scenario, the analysts are interested in the usage of synonyms. Can they spot any systematic replacement of words with synonyms from the first to the second text variant (T1 and T2)? First, they identify the correct category (3) in the menu and select it in the category bar. Then they consider the occurrences of synonym phenomena in the phenomenon bar and see that there are relatively few of them; clicking on a phenomenon rectangle focuses the annotation in the texts and allows to read the larger context of the respective passage, whereas right-clicking brings up the phenomenon detail window where they can see for example whether additional comments for a specific occurrence were made. Next, they choose the reduced text detail view. This way, almost all annotations in the text detail views can be seen at once (see Figure 8). Going through these passages, the analysts realize, that no systematic usage of synonyms can be found in this legend.

In the second scenario, the analysts consider syntactic rearrangements. Which passages of the first variant of a legend were rearranged in the second, where are they to be found, and is there a systematic co-occurrence with other phenomena (T1, T2, T3)? The analysts choose category 9 and first consider the text overview. They see that the phenomena of this category are spread all over the texts, but that they accumulate at the respective beginning and end. As they consider the details of the occurrences via right click on the phenomenon bar, they see that for the third occurrence, phenomenon 13c, also category 5 “changed semantics” was given – for this phenomenon, the detail window can be seen in Figure 2, the phenomenon bar cut-out in Figure 4, and one of text detail views in Figure 6. They choose category 5 additionally to see if there are more co-occurrences of these categories, however, as can be seen in the phenomenon bar, the phenomenon already found is the only one with an according color gradient. But there can be found two phenomena of category 5 and 9 which are situated close to each other in the phenomenon bar. Focusing them in the text detail views via clicking, the analysts see that the text segments of these phenomena overlap, and comparing these two examples, they may want to reconsider their annotation guidelines as to when the co-occurrence of 5 and 9 should be seen as a joint phenomenon and when as different phenomena with overlapping (or identical) text segments.

VI. EXPERT FEEDBACK AND DISCUSSION

The approach was discussed with and tested by the contributing medievalists in a series of about ten meetings. They considered it to be a valuable support for their analysis tasks, especially the filter features and the interlinked positions in all views were noted as helpful. They are interested in using the approach in future versions in their ongoing edition project of the HLR. Some of their suggestions for improvements have

already been implemented, and others are in the planning stage (see section VII).

The approach was also presented and discussed at a medieval studies workshop on the HLR (“Vorüberlegungen für eine digitale Edition und textgeschichtliche Untersuchung von *Der Heiligen Leben*, bearbeitete Redaktion (HL-Red. II)”, 08.02.2019, Stuttgart) and at a digital humanities workshop on digital editions in general (“Visualisierung in (digitalen) Editionen”, 21.05.2019, Düsseldorf). The general reception of the approach was good. However, one take-away from these events was that we need to sharpen the profile of the approach within the wider context of the edition project. It seemed not to be quite clear, whether the approach was aimed at the professional analyst who wants to explore the complex annotation data (and hence puts up with a steeper learning curve) or at a more casual, wider audience with less prior knowledge, who rather wants to read the two text variants. While our focus lays clearly on the former group, we think about introducing a “light version” to also accommodate the interest of the latter and to reconsider our design decisions in the light of this separation. Another point that spurred the discussion among the medievalist scholars was the elaboration of the phenomenon categories. As it was mentioned in section III, these categories will probably evolve further over the course of the edition project, which may also easily involve new visualization solutions.

VII. CONCLUSION AND FUTURE WORK

In this paper, we presented an approach to visually analyze the variance between two text versions which is accessible through a data set of rich, man-made text annotations of the phenomena of difference. It combines a set of coordinated close and distant reading views in an overview+detail design and provides means for filtering and detailed inspection of the annotation data. We grounded our design decisions in a task and requirement analysis that was conducted in close cooperation of the involved medievalists and visualization researchers. While our literature review yielded a rich variety of visualization approaches to the related problems of text reuse, text comparison and text collation, none of them would provide the functionality to support our specific task/data-set combination. The initial user feedback seems to support our conjecture that analysts may leverage the potential of this rich data set by employing our approach. We look forward to validate this result in a field study within the HLR edition project and to develop further approaches supporting the tasks therein. At the same time, while our approach originated from dealing with the specific text of the HLR, we believe that our lessons learned could easily be transferred to any use case with a similar data structure and task set.

As a first step in the further development of the approach after we had presented it on the occasions mentioned in the previous section, we implemented an editor functionality in order to allow the medievalists to annotate the texts directly within the tool and then export their annotated texts to an XML-file. This way, we can leverage our approach, whose

primary intention was to support the exploration of the annotation data, also to support the intricate task to create these data in the first place.

Besides the design variants already mentioned in section IV, which we would like to pursue as user options, we intend to direct our attention to the following aspects that are already inherent in the data:

- Show the manuscripts besides the transcriptions in the text views, together with the token- and annotation-information.
- Visualize the uncertainty indicators of the annotations; so far they are only shown in the phenomenon detail window.
- Right now, the approach juxtaposes the two text variants – experiment with superposition and/or explicit encoding designs instead; for a survey see Gleicher et al. [32].
- Right now, the overviews work on the level of a single legend. An overview visualizing larger portions of the whole HLR is a clear desideratum. For example, this could be achieved by means of visualizing some sort of fingerprint for each legend in the spirit of Jänicke et al. [15] or Keim and Oelke [33].

Then there are enhancements to the approach that would imply a more complex data structure and/or annotation process:

- A more complex, hierarchical system of phenomenon categories would call for a more sophisticated visualization concept.
- If the structure of the legends would be annotated as well – i.e. those segments that these formulaic texts quite typically are composed of, such as introduction, vita, miracle description, prayer etc. – this information could be visualized on an overview level as well.
- Further automatically derived data could be visualized, such as part of speech tags.
- If multiple annotators were to work on the same legends, the comparison of their results and an assessment of the inter annotator agreement could be supported by an appropriate visualization.

Finally, there are paths of development that would lead away from the current edition project. For example, the problem of visually comparing more than just two text variants (that are annotated for categorized differences), would probably call for a complete reworking of the design.

ACKNOWLEDGMENT

This work has been funded by the German Federal Ministry of Education and Research (BMBF) as part of the “CRETA” project. The prototype was implemented within the frame of a bachelor thesis that was performed by Wolfgang Knopki [34]. We would like to thank the three anonymous reviewers for their helpful comments and suggestions.

REFERENCES

- [1] K. Kunze, “Der Heiligen Leben,” in *Die deutsche Literatur des Mittelalters - Verfasserlexikon*, W. Stammer and K. Langosch, Eds. Berlin: de Gruyter, 1981, vol. 3, pp. 617–625.

- [2] M. Brand, K. Freenhagen-Baumgardt, R. Meyer, and W. Williams-Krapp, Eds., *Der Heiligen Leben: Band I: Der Sommerteil*, ser. Texte und Textgeschichte. Berlin: de Gruyter, 1996, vol. 44.
- [3] M. Brand, B. Jung, and W. Williams-Krapp, Eds., *Der Heiligen Leben: Band II: Der Winterteil*, ser. Texte und Textgeschichte. Berlin: de Gruyter, 2004, vol. 51.
- [4] W. Williams-Krapp, *Die deutschen und niederländischen Legendare des Mittelalters: Studien zu ihrer Überlieferungs-, Text- und Wirkungsgeschichte*, ser. Texte und Textgeschichte. Max Niemeyer, 1986, vol. 20.
- [5] S. Köbele, "Die Illusion der 'einfachen Form': Über das ästhetische und religiöse Risiko der Legende," vol. 134, no. 3, pp. 365–404, 2012.
- [6] J. Weitbrecht, M. Benz, A. Hammer, E. Koch, N. Nowakowski, S. Seidl, and J. Traulsen, *Legendarisches Erzählen: Optionen und Modelle in Spätantike und Mittelalter*, ser. Philologische Studien und Quellen, 2019, vol. 273.
- [7] A. Cockburn, A. Karlson, and B. B. Bederson, "A Review of Overview+Detail, Zooming, and Focus+Context Interfaces," *ACM Comput. Surv.*, vol. 41, no. 1, pp. 1–31, 2009.
- [8] F. Moretti, *Graphs, Maps, Trees: Abstract Models for a Literary History*. London: Verso, 2005.
- [9] S. Schreibman, "Versioning Machine 5.0: A Tool For Displaying & Comparing Different Versions of Literary Texts," <http://v-machine.org/>, 2016, 11.09.2019.
- [10] D. Wheelles and K. Jensen, "Juxta Commons," *J. Digit. Humanit.*, vol. 3, no. 1, 2014.
- [11] S. Koch, M. John, M. Wörner, A. Müller, and T. Ertl, "VarifocalReader: In-Depth Visual Analysis of Large Text Documents," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 1723–1732, 2014.
- [12] Z. Geng, T. Cheesman, R. S. Laramée, K. Flanagan, and S. Thiel, "ShakerVis: Visual Analysis of Segment Variation of German Translations of Shakespeare's Othello," *Inf. Vis.*, vol. 14, no. 4, pp. 273–288, 2013.
- [13] B. Asokarajan, R. Etemadpour, J. Abbas, S. Huskey, and C. Weaver, "Visualization of Latin Textual Variants using a Pixel-Based Text Analysis Tool," in *EuroVis Workshop Vis. Analytics*, 2016.
- [14] B. Asokarajan, R. Etemadpour, J. Abbas, S. Huskey, and C. Weaver, "TextTile: A Pixel-Based Focus+Context Tool For Analyzing Variants Across Multiple Text Scales," in *Eurographics Conf. Vis. (EuroVis) - Short Papers*, 2017, pp. 49–53.
- [15] S. Jänicke, A. Geßner, and G. Scheuermann, "A Distant Reading Visualization for Variant Graphs," in *Proc. Digit. Humanit.*, 2015.
- [16] S. Jänicke, A. Geßner, G. Franzini, M. Terras, S. Mahony, and G. Scheuermann, "TRAViz: A Visualization for Variant Graphs," *Digit. Scholarsh. Humanit.*, vol. 30, no. Suppl. 1, pp. i83–i99, 2015.
- [17] R. H. Dekker and G. Middell, "CollateX," <https://collatex.net/>, 2019, 13.09.2019.
- [18] T. Andrews, "The Stemmaweb Project," <https://stemmaweb.net/>, 2019, 13.09.2019.
- [19] M. Wattenberg and F. B. Viégas, "The Word Tree, an Interactive Visual Concordance," *IEEE Trans. Vis. Comput. Graph.*, vol. 14, no. 6, pp. 1221–1228, 2008.
- [20] P. Riehmann, H. Gruendl, M. Potthast, M. Trenkmann, B. Stein, and B. Froehlich, "WORDGRAPH: Keyword-in-Context Visualization for NETSPEAK's Wildcard Search," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 9, pp. 1411–1423, 2012.
- [21] S. Silvia, R. Etemadpour, J. Abbas, S. Huskey, and C. Weaver, "Visualizing Variation in Classical Text with Force Directed Storylines," in *VIS4DH*, 2016.
- [22] P. Riehmann, M. Potthast, B. Stein, and B. Froehlich, "Visual Assessment of Alleged Plagiarism Cases," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 61–70, 2015.
- [23] A. Abdul-Rahman, G. Roe, M. Olsen, C. Gladstone, R. Whaling, N. Cronk, R. Morrissey, and M. Chen, "Constructive Visual Analytics for Text Similarity Detection," *Comput. Graph. Forum*, vol. 36, no. 1, pp. 237–248, 2017.
- [24] S. Jänicke, T. Efer, M. Büchler, and G. Scheuermann, "Designing Close and Distant Reading Visualizations for Text Re-use," in *Computer Vision, Imaging and Computer Graphics - Theory and Applications*, S. Battiato, S. Coquillart, J. Pettré, R. S. Laramée, A. Kerren, and J. Braz, Eds. Cham: Springer, 2015, pp. 153–171.
- [25] M. Behrisch, M. Krstajic, T. Schreck, and D. A. Keim, "The News Auditor: Visual Exploration of Clusters of Stories," in *Int. Workshop Vis. Analyt. (EuroVA 2012)*, 2012.
- [26] S. Jänicke and D. J. Wrisley, "Visualizing Mouvvance: Toward a Visual Analysis of Variant Medieval Text Traditions," *Digit. Scholarsh. Humanit.*, vol. 32, no. Suppl. 2, pp. ii106–ii123, 2017.
- [27] S. Jänicke and D. J. Wrisley, "Interactive Visual Alignment of Medieval Text Versions," in *Proc. IEEE Conf. Vis. Anal. Sci. Technol. (VAST)*, 2017, pp. 127–138.
- [28] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann, "On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges," in *Eurographics Conf. Vis. (EuroVis) - STARS*, 2015, pp. 83–103.
- [29] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann, "Visual Text Analysis in Digital Humanities," *Comput. Graph. Forum*, vol. 36, no. 6, pp. 226–250, 2017.
- [30] M. Bostock, "D3: Data-Driven Documents," <https://d3js.org/>, 2019, 10.09.2019.
- [31] M. Harrower and C. A. Brewer, "ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps," *Cartogr. J.*, vol. 40, no. 1, pp. 27–37, 2003.
- [32] M. Gleicher, D. Albers Szafr, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, "Visual comparison for information visualization," *Inf. Vis.*, vol. 10, no. 4, pp. 289–309, 2011.
- [33] D. A. Keim and D. Oelke, "Literature Fingerprinting: A New Method for Visual Literary Analysis," in *Proc. IEEE Symp. Vis. Anal. Sci. Technol. (VAST)*, 2007, pp. 115–122.
- [34] W. Knopki, "Interaktive, vergleichende Visualisierung von Textveränderungen," Bachelor thesis, University of Stuttgart, Stuttgart, 27.09.2018.