



Journal of the Text Encoding Initiative

Issue 12 | July 2019 - May 2020

Selected Papers from the 2017 TEI Conference

Taking Note: Challenges of Dealing with Graphical Content in TEI

Martina Scholger



Electronic version

URL: <http://journals.openedition.org/jtei/2800>

DOI: 10.4000/jtei.2800

ISSN: 2162-5603

Publisher

TEI Consortium

Electronic reference

Martina Scholger, « Taking Note: Challenges of Dealing with Graphical Content in TEI », *Journal of the Text Encoding Initiative* [Online], Issue 12 | July 2019 - May 2020, Online since 29 January 2020, connection on 28 June 2020. URL : <http://journals.openedition.org/jtei/2800> ; DOI : <https://doi.org/10.4000/jtei.2800>

For this publication a Creative Commons Attribution 4.0 International license has been granted by the author(s) who retain full copyright.

Taking Note: Challenges of Dealing with Graphical Content in TEI

Martina Scholger

ABSTRACT

This paper describes the ongoing development of a TEI-conformant model for the encoding of hybrid primary sources containing text and graphical components on a similar level of semantic meaning. The subsequent considerations are part of the digital scholarly edition project representing the notebooks of the Austrian conceptual artist Hartmut Skerbisch (1945–2009), a handwritten corpus created over a period of almost 40 years. Besides text, drawings were an important visual means of expression for the artist. In the context of digitally representing the numerous graphics embedded in the notes with equivalent depth to the text, a three-step model is proposed. This model takes into account the (1) *graphical components* characterizing the composition, (2) *textual functions* describing and directing the execution of artworks, and (3) *interpretation* of the graphics and their contextualization with connected material and information. The paper discusses the existing methods of representing graphics in TEI, presents the combination of these methods in the actual scholarly edition project, and introduces the

semantic enrichment of the TEI sources through formal descriptions in RDF/XML taxonomies using linked open data. Thus, the genealogy of artistic concepts and artworks documented in the notebooks will become traceable.

INDEX

Keywords: art history, digital scholarly editions, notebooks, graphic encoding, text and image, genetic editing, Semantic Web

1. Introduction

- 1 The primary focus on text is not only inherent in the name of the Text Encoding Initiative itself, but also in the initial scope of the TEI as an encoding standard focusing on the notion of a text (Ide and Sperberg-McQueen 1995, 7–8; Burnard 2014, chap. 2). Undoubtedly, the TEI has become the means of choice for representing text in digital form in general and digital scholarly editing in particular. Although the TEI has always taken a generalistic approach, considering documents of any size, any number and any content (Burnard 2014), an extensive description of visual materials is nevertheless lacking in the TEI Guidelines. Since, in the wake of the large-scale digitization efforts of recent years, there is an increasing need to deal with hybrid source materials in digital humanities research, the TEI must also meet these challenges. Therefore, extended mechanisms and additions to the TEI Guidelines are required to open and expand the model for encoding non-textual sources with the same level of meaning as text.
- 2 Notes by artists, writers, and scientists are a good example for the necessity to accommodate hybrid written and pictorial sources. They are a valuable primary source of the story behind the scenes, tracing the origin, genesis, execution, and reuse of ideas, functioning as a storehouse of individual thoughts growing to mature concepts expressed in written and pictorial form (Radecke 2013, 150–151; Grésillon 1999, 11 and 168). The Dadaist writer Hugo Ball claimed that “the word and the image are one” (Ball 1974, 66), to emphasize the strong relationship between the acts of composing paintings and poetry. The Surrealist painter René Magritte proclaimed that “in a painting words are of the same substance as images” (Magritte 1929, 32). Consequently, in a notebook text and graphics enter into an inseparable dialogue.

- 3 This paper discusses a vivid example of such a hybrid source, namely an artist's notebook. In performative, ephemeral, and temporary art forms like conceptual art, performance art, and media art, notes play a key role in interpreting and understanding the actual artwork. They function as a documentary source, one of few remaining witnesses (or, in this case, the only one) of a temporally limited artistic event or artwork. This calls for a comprehensive representation of all related source components including notes and graphics.
- 4 This article will introduce a TEI-conformant approach to encoding sketches in the notebooks to a similar depth as the text itself. For this purpose, the TEI encoding of the actual record on a descriptive level is semantically enriched by factual statements in taxonomies using linked open data sources. The representation as a graph benefits the notebooks' inherent characteristics of distributed information by allowing for the implementation of flexible and reusable search mechanisms, the visualization of relationships both within the notebooks and outside, and the indexing of documents in different ways.

2. Project Context, Artist, and Source

- 5 The reflections on hybrid primary sources incorporating text and graphics on a similar semantic level and the challenges of digitally representing these phenomena discussed in this paper are part of an ongoing digital scholarly edition project on the notebooks by the Austrian conceptual and object artist Hartmut Skerbisch (1945–2009). It is a digital scholarly edition of the artist's notebooks with a special focus on the genesis of his artistic work. The overall goal of this project is to reconstruct the artist's inspirations powered by external influences and associative processes in the course of his developing individual manifestations: the evolution of a specific idea and its change over time.
- 6 Skerbisch expressed his ideas and concepts in various visual forms. In his early media artworks of the 1960s and 1970s, he examined in particular the space-generating properties of electronic media. In his work with photography starting in the late 1970s, he investigated the imaging process with its own means, i.e., addressing inherent qualities of photography such as seriality, indexicality, lack of materiality, and the production of reality in his own photographs (Frisinghelli 1994, 61–62). Similarly, his sculptural works from the late 1980s focused on the features of the material used and the recipients' spatial experience, which were more important to the artist than

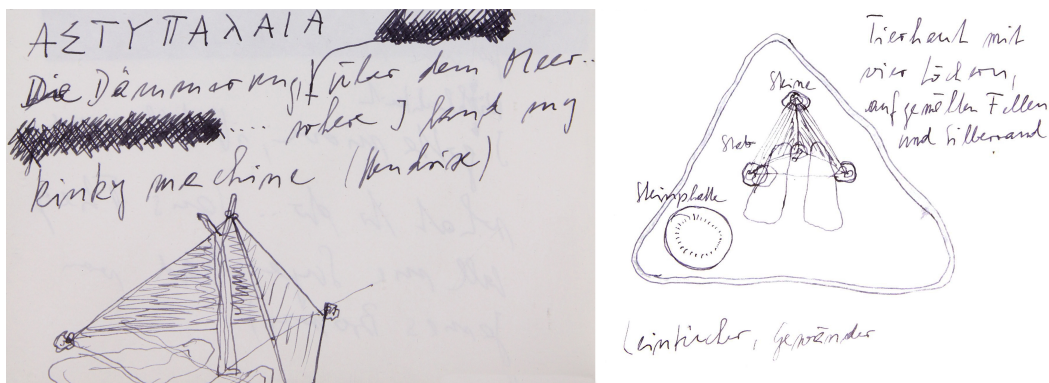
the aesthetic shell (Niegelhell 2015, 80). Skerbisch was trained as an architect and consequently he remained interested in the relationship between space, sculpture, and the spectator. This influence is recognizable in his clear constructional drawings and design language. He himself defined his work as a commentary on other works from literature, music, and art, a reflection on the likes of James Joyce, Franz Kafka, Kathy Acker, The Rolling Stones, or Jimi Hendrix (Verein der Freunde von Hartmut Skerbisch 2015). This circumstance explains the numerous references both in the texts and in the graphics. Skerbisch collected his ideas in a total of 35 notebooks, with approximately 2100 inscribed pages. These notebooks were created over a period of close to 40 years, between 1969 and 2008. The TEI-encoded notebook transcripts are stored, managed, and published in the digital repository GAMS of the Centre for Information Modelling at the University of Graz,¹ the Humanities' Asset Management System,² It is an OAIS-conformant long-term preservation system based on the open source project Fedora Repository³ (Steiner and Stigler 2018). A specially-designed content model stores the TEI encoding, related digital surrogates, RDF/XML statements extracted from the TEI during the upload into the repository, and dissemination methods operating on these datastreams in one single object. The current digital scholarly edition of Skerbisch's notebooks (Scholger 2018b) provides a synoptic view of facsimile and transcript, employing the IIF-based image viewer OpenSeadragon⁴ and a content-based view showing the references between single notebook entries of any kind and entities like persons, literature, music, artworks, and artistic concepts, which are presented in dedicated registers and thesauri. For the full context of the methodological considerations, the digital edition, and its technical implementation, see Scholger 2018a.

2.1 Importance of Visual Content

- 7 The TEI standard covers most of the textual phenomena contained in the notebooks: these range from easily-legible and interpretable continuous text to non-linear unstructured text fragments; single keywords where the content is difficult to interpret because of the lack of context; enumerations organized in columns; textual interventions like additions, deletions, substitutions, and transpositions; references to other areas on the page and metamarks as indicators of how the text should be read; and other source materials like patched newspaper clippings, glued-in photographs, and inserted slips of paper.

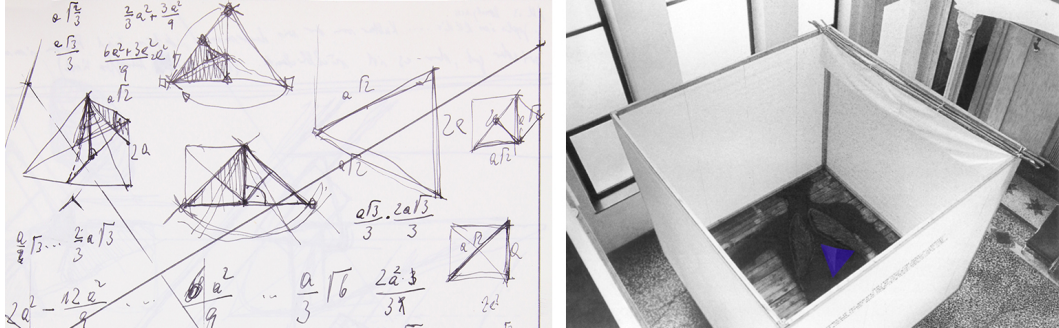
- 8 In addition to text, the notebooks contain a remarkable number of graphical representations, such as sketches, drawings, or scribbles (on approximately 450 out of 2100 pages). Skerbisch used this form of expression for recording his ideas and for preparing and planning works of art. Thus, they are merely accompanying the text instead of being illustrative, but they carry the same weight and equivalent meaning to the text itself. They can help to trace various stages of the realization of specific artworks and concepts as well to show developments and modifications of particular ideas.
- 9 **Figure 1** shows on the left a graphic with a view into a tetrahedron: its top is held up by a rod and an amorphous shape is recognizable on the bottom. It is actually a tent held by a branch, and the human indicated by a lung seeks shelter under the tent. The example on the right shows the same subject, but from an aerial perspective, and supplemented by a circle designated as a stone slab. The accompanying text gives information on the materials, colors, and objects to be used in the actual installation, and it labels specific components of the graphic.

Figure 1. Tent construction in two different views. Notebook 2: 2v and 10r.



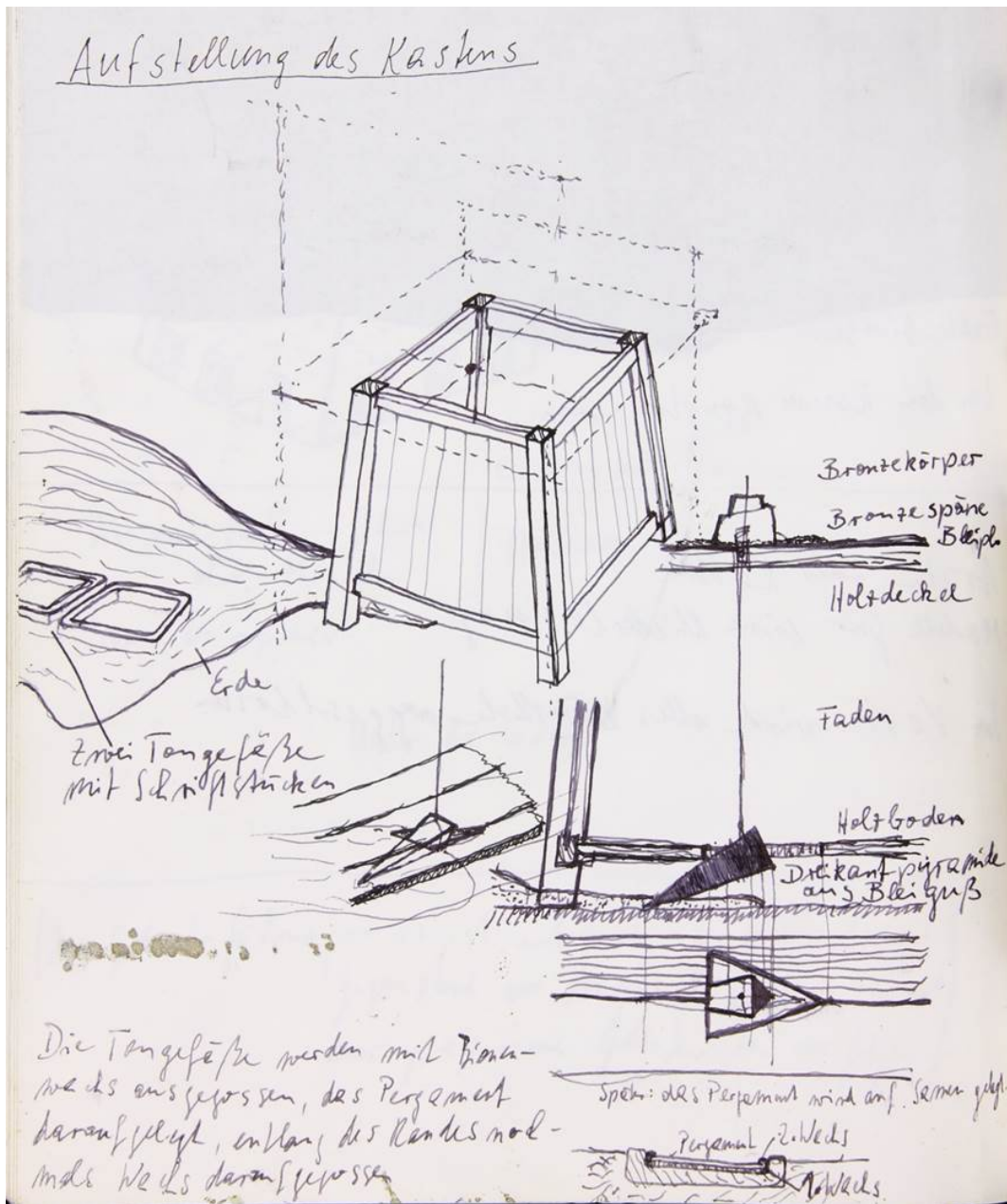
- 10 **Figure 2** shows another witness to the same topic: a notebook page with interlocking formulas, calculations, and constructional drawings with dimensions for the meticulous planning of a physical manifestation of the envisioned installation. Closer inspection of the triangular shapes reveals their relation to the tent construction. The photograph on the right shows one of the three known realized installations emerging from this design process: *Lispn!* from 1978 which features a tent construction.

Figure 2. Calculation of the tent construction (left), Notebook 4: 22r; Installation *Lispn!* (Photo: Hartmut Skerbisch) (right).



- 11 The complexity increases when different views and levels of details coincide on one page. Figure 3 shows a wooden box with an amorphous mass expanding on the bottom as a central representation in full view. Additionally, it shows four detailed renditions of individual areas of the full sketch from different angles. The drawings are accompanied by descriptive text and labels. Although the connection between the tent and the box is not given on a formal level, it becomes clear through the investigation of shared intellectual concepts, forms, and materials.

Figure 3. Constructional drawing with details and different views. Notebook 10: 15v.



- 12 These examples demonstrate the need to consider graphical representations as rich information carriers that require a specific description model. The central question is therefore how to develop a “notion of graphics,” i.e., how to appropriately treat graphical representations in a text-centric encoding environment.

2.2 Genealogy of Sketches

- 13 The transcription and digital representation of the textual entries in the notebooks mainly uses specific transcriptional elements from Chapter 11 “Representation of Primary Sources” of the TEI Guidelines (TEI Consortium 2019)⁵ to prepare the documentation of the origin and the genesis of the text. Drawing on the principles of genetic editing, mechanisms have been established by a dedicated working group on Genetic Editions focusing on the physical appearance and the writing process of a document rather than the intellectual structure (Burnard et al. 2010). This led to a thorough investigation of the French editorial school of the critique génétique (Grésillon 1999). While this movement focuses primarily on the creative writing process of literary manuscripts, Pierre-Marc de Biasi points out that the genetic approach can also be transferred to non-textual material, especially to architecture and sculpture (De Biasi 1997, 124). In the case of the present notebooks used for the conception of artworks, the constitution of the text plays a subordinate role to the unfolding of artistic ideas and general concepts. According to this assertion, the genetic approach is applied to graphical representations. Sketches, formulas, and other graphical representations are integral forms of artistic expressions. Drawings can carry a specific message which cannot always be expressed with words. Here, the graphic becomes the primary object, and the text moves into the background. Representing visual material like graphics is thus not the task of—nor suitably dealt with by—text-image-linking tools, for which a range of satisfying solutions already exist (OpenSeadragon,⁶ Mirador,⁷ DFG-Viewer,⁸ etc.). Instead, it is necessary to accurately describe and encode the type and design of graphical representations as well as the subservient functions of text in a formal, comparative, and processable way. How then can the researcher apply the genetic approach to sketches and graphical elements in the TEI context? The sketches in figure 4 exemplify the development between the various graphical representations, which is in fact comparable to textual alterations in genetic criticism.

Figure 4. Development stages of three drawings of a tent. Notebook 4: 19v.

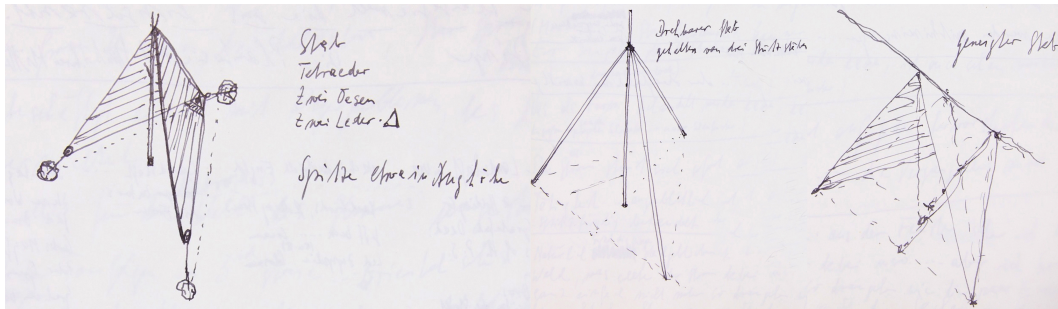


Figure 4 shows three stages of a tent construction from an axonometric view. The most significant change can be seen in the type of the supporting structure: the first construction shows a single rod, placed in the middle, with the textile spread from its top; the second one shows a rod placed in the middle with three rods attached to it, forming a tetrahedron; in the third construction the centre of gravity is shifted and the rod is tilted. These different stages have to be formally described.

3. Representing Graphics with TEI

- 14 In the following, the current possibilities and their pitfalls for representing graphics in TEI are discussed. Resulting from this consideration, a TEI-conformant solution for a detailed and formal description of graphical representations in a text-centric encoding environment is presented. Several approaches to encoding photographs, illustrations, graphics, and other visual material that can appear as part of a text are already featured in the TEI Guidelines in chapter 14.4. on “Specific Elements for Graphic Images” (TEI Consortium 2019)⁹ and 11.1 on “Digital Facsimiles” (TEI Consortium 2019).¹⁰ The simplest method, providing a mere indication of a graphic’s existence embedded in a continuous text, is recorded by the element `<graphic>` with a reference to the location of the digital surrogate in the `@url` attribute in form of a URI-reference, as shown in [example 1](#):

Example 1. Recording a simple graphic.

```
<graphic url="tent.jpg" />
```

- 15 This simple mechanism is useful in the context of illustrations accompanying the text, where the image has illustrative character and not an explanatory function. This minimal encoding, registering the graphic's existence, can be expanded through nesting the graphic reference `<graphic>`, the prose description `<figDesc>`, and the textual transcription in arbitrary text blocks `<ab>` within a `<figure>` element employing a structure-oriented encoding approach. [Example 2](#) demonstrates the use of this expansion for the first drawing of [figure 4](#).

Example 2. Encoding of graphic and text.

```
<figure>
  <graphic url="tent.jpg" />
  <figDesc>Zeltkonstruktion mit zentralem Zwieselstab [tent construction with
  a central rod]</figDesc>
  <ab>Stab (rod) <lb/>Tetraeder [tetrahedron] <lb/>Zwei Oesen (two eyelets)
  <lb/>Zwei Leder  $\Delta$  [two leather triangles]</ab>
  <ab>Spitze etwa in Augenhöhe [peak near eye level]</ab>
</figure>
```

- 16 When focusing on the physical disposition of the document and the writing (and drawing) process rather than the textual structure, the `<sourceDoc>` element has to be considered. [Example 3](#) uses the same source and shows the segmentation of textual and visual content into surfaces, zones, and lines with exact coordinates for locating every discrete graphical component and portion of text.

Example 3. Encoding of the physical disposition of graphics and their components.

```
<sourceDoc>
  <surface>
    <zone ulx="288" uLy="136" lrx="1500" lry="892">
      <graphic url="tent.jpg" />
      <zone ulx="947" uLy="193" lrx="1083" lry="257">Stab</zone>
      <zone ulx="947" uLy="193" lrx="1083" lry="257">Tetraeder</zone>
      <zone ulx="944" uLy="325" lrx="1147" lry="378">Zwei Oesen</zone>
      <zone ulx="942" uLy="378" lrx="1206" lry="435">Zwei Leder  $\Delta$ </zone>
      <zone ulx="936" uLy="479" lrx="1466" lry="597">Spitze etwa in
      Augenhöhe</zone>
    </zone>
  </surface>
</sourceDoc>
```


- 17 But neither of these options carries the necessary complexity to model the graphics on the same formal and content-related level as the text in the notebooks. Using these approaches, it is hardly possible to describe their design and meanings, to formalize the complex interactions of text and graphical components, and to record alterations and modifications for comparing various draft stages.
- 18 An attempt to assign meaning to visual material in TEI-encoded artifacts was made by Martin Holmes's *Image Markup Tool* which provides image annotation within the <body> and corresponds with dedicated zones of the image in the <facsimile> section which in turn refers to descriptive categories stored in the encoding description of the TEI Header (Holmes 2012). Building upon this approach in order to study the genesis of artworks documented in the notebooks, a formalized description in combination with semantic web methods facilitates the comparison of the preliminary drawings and the recording of developments as well as their contextualization with regard to other notebook entries (text and graphics), overall intellectual concepts, and actual artworks.
- 19 For a comprehensive description of graphical representations, a multi-layered model¹¹ is proposed, accounting for the various (1) graphical components, (2) textual functions, and (3) editorial interpretations. The first two levels, describing the graphical components and the textual functions—essentially constituting the material record of the source—are descriptive. The third or interpretational layer allows for commenting and contextualizing the material with links to internal and external entities, and is provided by the editor. This refers to the editorial dichotomy of Hans Zeller in separating the objective facts—the *Befund* (record)—and its further editing—the *Deutung* (interpretation)—of the record (Zeller 1971, 50–52). This methodological differentiation of editorial tasks in philology is close to Erwin Panofsky's theory on the description and interpretation of works of art: his three-level model consists of a *pre-iconographic analysis*, describing the formal aspects of the work in question; the *iconographic analysis* deals with the interpretation; and the *iconological analysis* considers the artwork in its particular cultural and historical context (Panofsky 2006, 37–40).
- 20 The first layer of the proposed model describes—independently of a specific technical implementation—the graphical components of a figure. It declares a) the *type* of the graphic representation (e.g., sketch, constructional drawing, doodle), b) the *projection* (e.g., front view, plan

- view), c) the *status* of execution (e.g., total view, detail view), d) the material of the *information carrier* (e.g., paper, newspaper, photograph), e) the *drawing instrument* (e.g., pencil, ink pen), f) the *date* or *time* span (to facilitate a chronological order for further investigation on the genesis of the work), and g) the graphical *shapes and figures* primarily used (e.g., triangle, square, cube, tetrahedron).
- 21 The second layer records the textual functions. It refers to any explanatory text accompanying the graphic provided by the artist for descriptive purposes. This category includes a) a *caption*, b) a *description* related to the whole graphic or parts of it, and c) a *label* which designates a specific component of the graphic, sometimes made explicit through a connecting line or clarified through its distinct positioning. Additionally, the textual content can be of a specific type (e.g., providing information on the material or measurements to be used in the physical manifestation).
- 22 The third layer represents the editor's interpretation. The model consists of a) a *general comment*, b) several *relations*—to other *notebook entries*, to *external entities* from literature, art, music, or other preparatory objects (like physical models or photographs), to actual *manifestations*, and to general theoretical *concepts* reflected on by the artist—and c) the *image genesis*, which is not explicitly encoded but is a result of the indexing process and therefore operative: variants are generated automatically and can be brought into a sequence to show—in a manner comparable to different stages of a text in genetic criticism—the development of modifications. The factual information is represented as RDF statements in dedicated taxonomies, enriched with information from existing authority files where applicable.
- 23 The model in its reduced form neglects those categories that are either already applied as general metadata to the notebook (e.g., author, archive, copyright information), or those that do not apply to Skerbisch's graphics. If required, the model can easily be extended.
- 24 The application of these categories to the tent in figure 4 is shown in tables 1, 2, and 3. Constants and alterations are easily identified and documented. Most of the properties—like type, projection, shapes, material, notebook entry, external reference, manifestation, and concept—refer to thesauri and controlled vocabularies from existing authority files, especially the Art & Architecture Thesaurus (AAT)¹² and the individually compiled catalogue of Skerbisch's artworks¹³ and artistic concepts.¹⁴ Tables 1, 2, and 3 reveal that many of the instances remain the same across

the three representations, but some show distinct differences, especially when it comes to the textual functions which contain instructions for materiality and measurements for the physical realization of the artwork.

Table 1. Instances of the graphical components.

	graphic 1	graphic 2	graphic 3
type	sketch		
projection	axonometry		
status	total view		
information carrier	paper		
drawing instrument	ink pen, black		
date	1973–08 (not before)		
form/shape	triangle, tetrahedron		

Table 2. Instances of the textual functions.

	graphic 1	graphic 2	graphic 3
caption	Studien zum Haus am Lepenen Strudel [studies on the house at the Lepenen Strudel]		
description	Spitze etwa in Augenhöhe [peak near eye level]	Drehbarer Stab gehalten von drei Stützstäben [rotating rod held by three support rods]	-

label	Stab; Tetraeder; Zwei Oesen; Zwei Leder Dreiecke [rod; tetrahedron; two eyelets; two leather triangles]	-	Geneigter Stab [tilted rod]
measure	~1,5-1,7m	-	-
material	Leder [leather]		

Table 3. Instances of the interpretational layer.

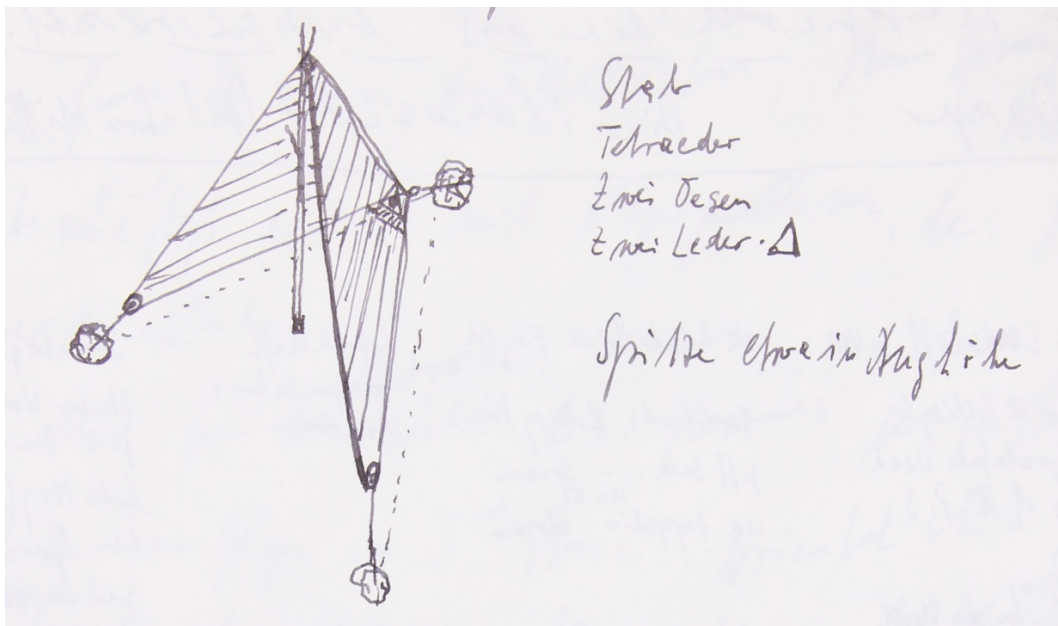
	graphic 1	graphic 2	graphic 3
comment	tent construction with a central rod	tent construction with a rotating rod set up in the middle	tent construction with a tilted rod at the rear edge
Relation			
notebook entry	graphic (#TB01-004); text (#TB04-128); text (#TB05-12) etc.		
external reference	photograph (P112)		
manifestation	Erde. Our cubehouse still rocks (A10003); Lispn! (A10036); reden blattartig (A10004)		
concept	body, environment, settlement		
Image genesis			
alteration	central rod	rotating central rod	one tilted rod as supporting element

addition		three additional bracing rods	side panels added
deletion		side panels removed	

3.1 Proposal for an Encoding Model

- 25 To transfer the demonstrated model into digital structures, the existing methods for encoding graphics in TEI—already introduced in the paragraph on *Representing Graphics with TEI*—are brought together and enhanced with a semantic web approach. The proposed means of encoding the graphic from figure 5 in TEI is as follows:

Figure 5. Sketch representing a tent, with additional textual functions. Notebook 4: 19v.



Example 4. Alignment of transcription and facsimile.

```

<!-- section 1: encoding of the physical dispositions in the facsimile structure
-->
<facsimile>
  <surface xml:id="fol_19v">
    <zone xml:id="F-27" ulx="316" uly="145" lrx="1482" lry="837">
      <graphic url="tent.jpg"/>
    
```

```

<zone xml:id="F-27-01-a" ulx="947" uly="193" lrx="1083" lry="257"/>
<zone xml:id="F-27-01-b" points="595,515 610,194 620,193 615,514"/>
<zone xml:id="F-27-02-a" ulx="947" uly="267" lrx="1143" lry="317"/>
<zone xml:id="F-27-02-b" points="340,561 615,155 820,345 699,847 631,423"/>
<!-- more <zone> elements -->
</zone>
</surface>
</facsimile>
<!-- section 2: encoding of the sketch in the intellectual text structure -->
<figure facs="#F-27" ana="art:A10003 gt:1010100 gt:6070000 gt:7010000">

<figDesc>Zelt mit Zwieselstab</figDesc>

<figure facs="#F-27-01">
  <label facs="#F-27-01-a" ana="gt:7050000">Stab</label>
  <graphic url="#F-27-01-b"/>
</figure>

<figure facs="#F-27-02">
  <label facs="#F-27-02-a" ana="gt:5030000">Tetraeder</label>
  <graphic url="#F-27-02-b"/>
</figure>

<figure facs="#F-27-03">
  <label facs="#F-27-03-a" ana="gt:2080000">Zwei Oesen</label>
  <graphic url="#F-27-03-b"/>
</figure>

<figure facs="#F-27-04">
  <label facs="#F-27-04-a">Zwei <material ana="gt:2020000">Leder</material> Δ</
label>
  <graphic url="#F-27-04-b"/>
</figure>

<graphic url="#F-27-a"/>
<label facs="#F-27-b">Spitze etwa in <material type="height"
ana="gt:6030000">Augenhöhe</material></label>

```

</figure>

The encoding in [example 4](#) consists of two parts, aligning the transcription with the facsimile, using the parallel transcription approach. For representing the respective detail of the digital surrogate and the physical disposition of the graphic, the facsimile structure, i.e., <facsimile>, <surface>, <graphic>, and nested <zone> elements, is used. In contrast to the previously presented method for representing the digital surrogate and the physical disposition within <sourceDoc>, which per definition is used for a single source document (TEI Consortium 2019), the <facsimile> structure is employed. Coordinates using members from the attribute class @att.coordinated locate not only the graphic as a whole but also the position of single components, both graphical components and textual functions. This supports the orientation within the graphic and the comparability of details when further processing the content, for example for presentation in an image viewer. A number of tools are available to facilitate the determination of the coordinates. For marking rectangular areas, the [TEI-Facsimile-Plugin](#)¹⁵ as part of the oXygen XML Editor is convenient since this occurs in the same working environment as used for the encoding, inserting the code at the intended place. Unfortunately, the plugin does not support polygons. For this purpose, the [TEI-Zoner](#),¹⁶ a simple browser-based tool developed by Chris Sparks for marking single points on a surface, is employed. The resulting TEI-conformant snippet—<zone> elements representing the coordinates as pairs of values in the @points attribute—can be copied directly to the corresponding position within the facsimile structure in the actual TEI document. For further elaboration on connecting lines between labels and specific graphic details, the <path> element for non-closed areas is available, recently introduced with version 3.4.0 of the TEI Guidelines (TEI Consortium 2019, <path>).¹⁷ For encoding a graphical representation within the intellectual text structure in the body, the <figure> element is used, nesting the <figDesc> element for a summarizing prose description of the setting provided by the editor. The TEI-schema allows for nested <figure> elements in order to localize specific parts of the graphic—represented each time through a nested <graphic> element—and to localize and encode the textual function and instruction which is represented through a <label> element. In the TEI Guidelines, <label> is defined as containing “any label or heading used to identify part of a text, typically but not exclusively in a list or glossary” (TEI Consortium 2019, chap.

`<label>`).¹⁸ This supplementary phrase allows for a broader interpretation and the use of the `<label>` element in this context. In case the artist gave additional instructions concerning measurements and materials for the planned installation, they are encoded using the elements `<measure>` and `<material>`, although the latter was originally introduced for describing the physical material of manuscripts (TEI Consortium 2019, chap. `<material>`).¹⁹ The relation between the intellectual content and the physical disposition is established by the `@facs` attributes on the figure tagset, pointing to the designated `@xml:id` attributes on the corresponding `<zone>` elements. This largely covers the description of the text functions, but the formalization of the graphic-specific properties is still missing. Descriptive elements from the TEI/XML encoding are semantically enriched and formalized by linking them to dedicated statements on subjects in taxonomies expressed in RDF/XML using linked open data sources. This is implemented using the `@ana` attribute which “indicates one or more elements containing interpretations of the element on which the `@ana` attribute appears” (TEI Consortium 2019).²⁰ This trinity of image, sign, and meaning of content (Vogeler 2015) facilitates a) the formalized representation of the TEI-encoded content, b) the retrieval of common factual statements, and c) the visualization of relations between different entities. The next step will be to formalize a special case of the `<figure>` element in a project-specific ODD in order to further implement the proposed model. In example 4, the value “art:A10003” refers, indicated through its prefix “art”, to an RDF triple of a description from the artwork thesaurus, which contains all known artworks by Skerbisch, with related information. The prefix “gt” addresses subjects from the graphic thesaurus, which incorporate concepts from existing authority files in addition to individually defined concepts where the authority files are lacking. It can be flexibly expanded at any time when new material is edited. The graphic thesaurus represented in SKOS currently contains seven top concepts: 1) visual works, 2) materials, 3) drafting, drawing and writing equipment, 4) supporting material, 5) geometric figures, 6) views, and 7) interpretation. In summary, this approach of a) localizing the topographic aspects of the graphic in the `<facsimile>` section, b) graphic description through `<figure>` elements embedded in the text structure, and c) semantic enrichment using the `@ana` attribute allows for the detection and extraction of graphics related by their content throughout the entire notebook corpus.

Comparing the changes with a formal model clearly shows the alteration of key components and concepts, i.e., *image genesis*. Hence, the encoding of related concepts allows the extraction, comparison, and genealogical sequencing of related sketches from throughout the corpus.

- 26 Similar concepts assigned to graphics can be retrieved through SPARQL queries: during the upload process into the repository, information encoded in the TEI document is formalized as RDF triples which are stored in a triple store for further retrieval.
- 27 [Example 5](#) shows the RDF triple with the actual statements representing a single graphical representation (see [figure 5](#)) from the notebook, indicated as a URI in the `@rdf:about` of the `<rdf:Description>`. The reference to the artwork thesaurus, an RDF/XML file, is determined by the predicate `<g2o:isArtworkID>`; the predicate `<rdf:type>` references concepts from the graphic thesaurus via a URI in the `@rdf:resource`.

Example 5. Extracted RDF triples on a single graphical representation.

```
<rdf:Description rdf:about="http://gams.uni-graz.at/o:sker.04#GRA-046-0">
  <dc:type rdf:resource="http://purl.org/dc/dcmitype/Image" />
  <g2o:isArtworkID rdf:resource="http://gams.uni-graz.at/o:sker.artworks#A10003"/>
</rdf:Description>
<dc:date>1972</dc:date>
<dc:title>Erde (Our cubehouse still rocks)</dc:title>
<rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#1010100"/>
  <rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#2020000"/>
  <rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#2080000"/>
  <rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#5030000"/>
  <rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#6030000"/>
  <rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#6070000"/>
  <rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#7010000"/>
  <rdf:type rdf:resource="http://gams.uni-graz.at/skos/scheme/o:sker.graphics/
#7050000"/>
```



```

<g2o:image rdf:resource="http://gams.uni-graz.at/iiif/o:sker.04/
IMAGE.046/288,136,1212,756/pct:20/0/default.jpg"/>
<g2o:part0f rdf:resource="http://gams.uni-graz.at/o:sker.04"/>
</rdf:Description>

```

The storage of the RDF triples extracted from the TEI document and all related thesauri and taxonomies in the triple store Blazegraph lays the foundation for a common query of the data and the retrieval of shared artistic concepts. The following SPARQL query in [example 6](#) first finds all notebook fragments (subject), where the predicate `g2o:isArtworkID` refers to a specific concept (object) in the artwork thesaurus (`art:A10003`) and extracts the artwork title via the predicate `dc:title`. The following graph pattern restricts the result with the predicate `g2o:image` to all image fragments and delivers the image coordinates as a URI for later presenting the designated detail through the [IIIF Image API](#) using image request parameters.²¹ It joins together triples matching any number of the `rdf:type` predicate referencing concepts from the graphic thesaurus: sketches (`gt:1010100`), skin (`gt:2020000`), eyelets (`gt:2080000`), tetrahedra (`gt:5030000`), eye-level-views (`gt:6030000`), axonometric projection (`gt:6070000`), tents (`gt:7010000`), and rods (`gt:7050000`). The predicate `skos:prefLabel` returns the label of the concept from the SKOS thesaurus and `dc:date` extracts the closest date information—either from the pertinent entry, from the notebook itself, or from the creation date of related artworks—to the graphic detail.

Example 6. SPARQL query to retrieve notebook fragments with specific properties.

```

PREFIX dc:<http://purl.org/dc/elements/1.1/>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
PREFIX g2o:<http://gams.uni-graz.at/ontology#>
PREFIX gt:<http://gams.uni-graz.at/skos/scheme/o:sker.graphics/#>
PREFIX art:<http://gams.uni-graz.at/o:sker.artworks#>

select * where

{
  ?Fragment g2o:isArtworkID art:A10003 .
  art:A10003 dc:title ?artworkTitle .
}

```

```
{
  ?Fragment g2o:image ?imageRef .
  ?Fragment rdf:type gt:1010100 .
  ?Fragment rdf:type gt:7010000 .

  {gt:1010100 skos:prefLabel ?label .} UNION
  {gt:7010000 skos:prefLabel ?label .}

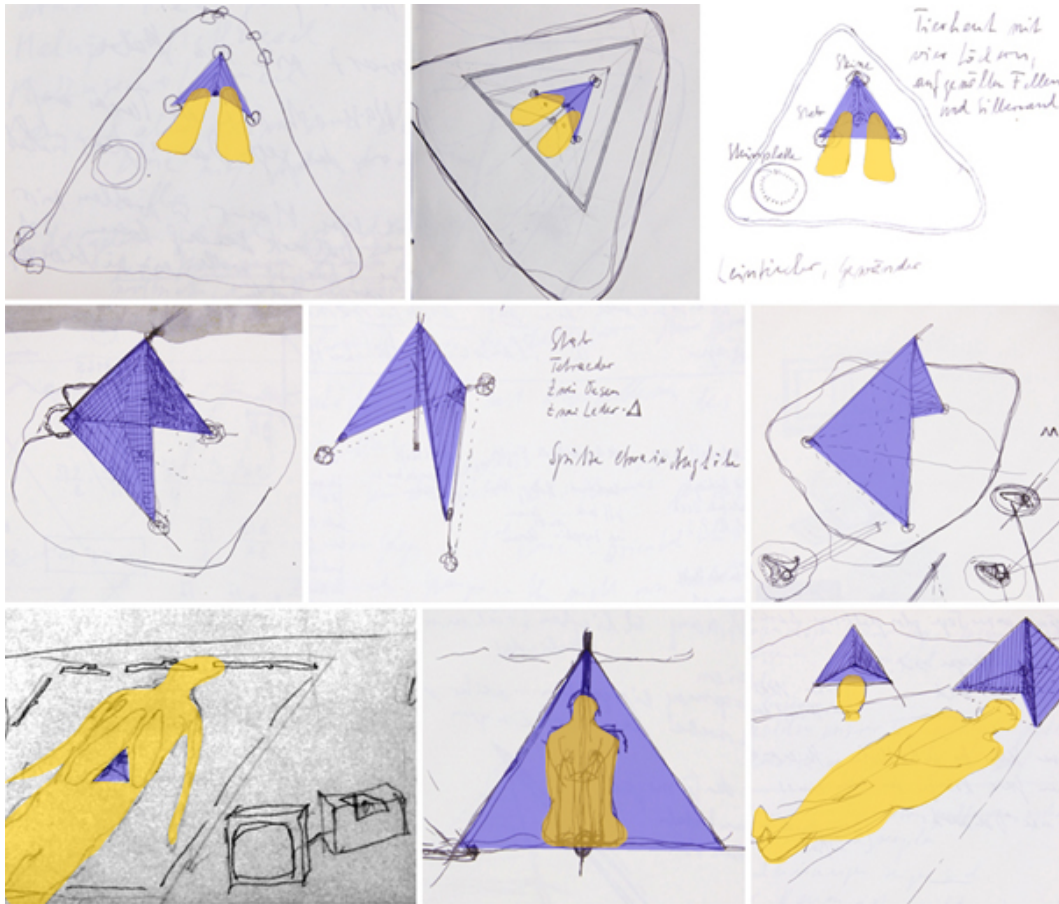
  FILTER (lang(?label) = 'eng')
}

OPTIONAL {
  ?Fragment dc:date ?date .
}

}
```

This SPARQL query briefly exemplifies the flexibility of querying data with different restrictions and variations to address the main research question: the genesis of concepts and artworks throughout the notebook corpus. Through this formalization process, similar graphics are extracted, compared, and examined for similarities and differences. The resulting XML data from this SPARQL query returns a list of images as well as a synopsis of features. The reference to the position in the notebooks gives relative temporal information to create a chronology of changes. This is implemented as a search function, dynamically transforming the results to a web representation. Drawing on the encoding of the location of graphical details explained above, single components are highlighted using JavaScript and thus support the interpretation process (figure 6).

Figure 6. Highlighting of similar properties in different sketches.



- 28 The comparative juxtaposition, in combination with the temporal dimension of the course of development, reveals specific stages of development and degrees of completion. The SPARQL query extracts, in a first step, all sketches showing a tent (gt:7010000), highlighted in blue. A further restriction on the subject human figures (gt:7020000), highlighted in yellow, shows only graphics with a combination of tents and human representations. Finally, adding an axonometric projection (gt:6070000) as a further selection criterion leaves only the left and the right graphics of the last row. The encoding of shapes on the <zone> element in [example 4](#) is relatively simple, but, depending on the level of detail and the complexity of the shape itself, the process can be difficult and time-consuming.

- 29 This perspective on the corpus draws a picture of the chronological sequence in which the graphics occurred and shows the intensity with which the artist has devoted himself to a specific topic in the conception of his artworks. Since the notebook entries are also linked to artworks and artistic concepts, this method allows searches for corresponding text passages in addition to the graphics. Through the combination of textual entries, graphical representations, external links, and work manifestations, a comprehensive analysis of the work creation process becomes possible.

4. Closing Remarks and Future Developments

- 30 A quotation from Hartmut Skerbisch himself aptly summarizes the challenge at hand: “Though not self-evident, there are after all also works of art. Once we get involved, we are in a labyrinth of relationships and seek a way out” (Skerbisch 2000).
- 31 Skerbisch’s artworks usually evolved over a long period of time. This becomes evident in the notebooks themselves: various text fragments and graphical representations are scattered throughout the corpus, and ideas undergo numerous variations. Based on the different manifestations shown in [figure 7](#), it is a challenge to understand the ideas behind them without familiarity with the artist. The tent appears in a private photograph from 1969 and is taken up again in the artworks *Erde. Our cubehouse still rocks* from 1976 and yet again two years later in the installation *Lispn!*. The latter refers to a quotation from James Joyce’s *Finnegans Wake* “Soft morning city! Lsp! I am leafy speafing Lpfl ... Lispn!” (Joyce 1947, 619) which also reiterates *reden blattartig*, a 1976 literal translation of *leafy speafing* into German. All of these artworks share the concepts of “breathing” and “body” (incorporating the shape of a human lung) and “settlement” (expressed through the tent).
- 32 Investigating the graphical representations in notebooks—and especially their changes and alterations in material, positions, components, proportions, etc.—can assist us in solving the bigger puzzle of reconstructing an artist’s creative processes. Such investigations should therefore be much more prominent in the genre of digital scholarly editing. To achieve that goal, however, it is essential to encode graphics with a similar depth and complexity to text. For this purpose, a combination of TEI encoding and semantic web technologies seems to be a promising solution.

Figure 7. Evidence of graphics related to the same concept. *reden blattartig*, 1976; *Erde. Our cubehouse still rocks*, 1976; *Lispen!*, 1978. (Photos: Hartmut Skerbisch).



- 33 The proposed encoding model for graphical representations is currently a prototype, which is applied to the whole notebook corpus. Additionally, its suitability to comparable source material which shares similar characteristics is yet to be investigated. One can imagine that notebooks by scientists (e.g., laboratory books by Alexander Fleming),²² artists (e.g., sketches and notes by Albrecht Dürer),²³ scientific travel correspondence (e.g., Alexander von Humboldt),²⁴ notes from architects, engineers, and even medieval manuscripts will benefit from the application of the proposed model. Only through the combination of an equally thorough encoding of textual and graphical components, and their semantic enrichment, can such hybrid source material be fully understood.

BIBLIOGRAPHY

- Ball, Hugo. 1974. *Flight Out of Time: A Dada Diary*. The documents of 20th-century art. New York: Viking Press.
- Burnard, Lou. 2014. *What is the Text Encoding Initiative?* OpenEdition Press. doi: 10.4000/books.oep.426.
- Burnard, Lou, Fotis Jannidis, Elena Pierazzo, and Malte Rehbein. 2010. "An Encoding Model for Genetic Editions." <http://www.tei-c.org/Vault/TC/tcw19.html>.
- De Biasi, Pierre-Marc. 1997. "Horizons for genetic studies." *Word & Image* 13 (2): 124–34. doi: 10.1080/02666286.1997.10434277.
- Frisinghelli, Christine. 1994. "Der wesentliche Bestand der Erscheinungen." In *Hartmut Skerbisch: Werkauswahl 1969–1994, Neue Galerie Graz*, edited by Werner Fenz, 60–65. Graz: Neue Galerie.
- Grésillon, Almuth. 1999. *Literarische Handschriften: Einführung in die "critique génétique"*. Arbeiten zur Editionswissenschaft 4. Bern: Lang.

- Holmes, Martin. 2012. "Image Markup Tool." Accessed May 25, 2018. https://hcmc.uvic.ca/~mholmes/image_markup/.
- Ide, Nancy M., and C. M. Sperberg-McQueen. 1995. "The TEI: History, Goals, and Future." *Computers and the Humanities* 29 (1): 5–15. <https://doi.org/10.1007/BF01830313>.
- Joyce, James. 1947. *Finnegans Wake*. New York: Viking Press.
- Magritte, René. 1929. "Les mots et les images." *La Révolution Surréaliste* 12: 32–33.
- Niegelhell, Franz. 2015. "Der Zeit Voraus Sein." In *Hartmut Skerbisch: Leben Und Werk, Life and Work: Gegenwart Als Gegenwart, Present as Present*, edited by Verein der Freunde von Hartmut Skerbisch, 20–97. Wien: Verlag für Moderne Kunst.
- Panofsky, Erwin. 2006. *Ikonographie und Ikonologie: Bildinterpretation nach dem Dreistufenmodell*. Köln: DuMont.
- Radecke, Gabriele. 2013. "Notizbuch-Editionen: Zum philologischen Konzept der Genetisch-kritischen und kommentierten Hybrid-Edition von Theodor Fontanes Notizbüchern." In *Internationales Jahrbuch für Editionswissenschaft*. Vol. 27, edited by Rüdiger Nutt-Kofoth and Bodo Plachta. 149–72. Münster: De Gruyter.
- Scholger, Martina. 2018a. "Assoziationsprozessen auf der Spur: Digitale Edition der Notizbücher von Hartmut Skerbisch." PhD diss., University of Graz. <http://unipub.uni-graz.at/urn:nbn:at:at-ubg:1-123592>.
- . 2018b. *Die Notizbücher von Hartmut Skerbisch. 1969–2008*. Eine digitale Edition. Accessed May 25, 2018. <https://gams.uni-graz.at/skerbisch>.
- . "Pieces of a Bigger Puzzle: Tracing the Evolution of Artworks and Conceptual Ideas in Artists' Notebooks." In *Versioning Cultural Objects. Digital Approaches*, edited by Roman Bleier and Sean M. Winslow. 27–58.
- Skerbisch, Hartmut. 2000. "Statements." Accessed May 25, 2018. <http://www.van.at/alt/van99/energie/statement.htm>.
- Steiner, Elisabeth, and Johannes Stigler. 2018. "GAMS—An infrastructure for the long-term preservation and publication of research data from the Humanities." In *Mitteilungen der VÖB 71: Nr. 1: Repositorien in Österreich*, edited by Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 66–75. doi: 10.31263/voebm.v71i1.1992.
- TEI Consortium. 2019. *TEI P5: Guidelines for Electronic Text Encoding and Interchange. Version 3.5.0*. <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/>.
- Verein der Freunde von Hartmut Skerbisch, ed. 2015. *Hartmut Skerbisch: Leben Und Werk, Life and Work: Gegenwart Als Gegenwart, Present as Present*. Wien: Verlag für Moderne Kunst.

- Vogeler, Georg. 2015. "Warum werden mittelalterliche und frühneuzeitliche Rechnungsbücher eigentlich nicht digital editiert?" In *Grenzen und Möglichkeiten der Digital Humanities*, edited by Constanze Baum and Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. http://dx.doi.org/10.17175/sb001_007.
- Zeller, Hans. 1971. "Befund und Deutung: Interpretation und Dokumentation als Ziel und Methode der Edition." In *Texte und Varianten: Probleme ihrer Edition und Interpretation*, edited by Gunter Martens and Hans Zeller, 45–89. München: Beck.

NOTES

- 1 Centre for Information Modelling – Austrian Centre for Digital Humanities, accessed February 8, 2019, <https://informationsmodellierung.uni-graz.at/en/>.
- 2 Humanities' Asset Management System GAMS, accessed May 25, 2018, <https://gams.uni-graz.at/>.
- 3 Fedora Repository (Flexible Extensible Digital Object Repository Architecture), accessed February 8, 2019, <http://duraspace.org/fedora/>.
- 4 OpenSeadragon, accessed February 8, 2019, <https://openseadragon.github.io/>.
- 5 <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/PH.html>.
- 6 OpenSeadragon, accessed February 8, 2019, <https://openseadragon.github.io/>.
- 7 Mirador, accessed February 8, 2019, <http://projectmirador.org/>.
- 8 DFG-Viewer, accessed February 8, 2019, <https://dfg-viewer.de/>.
- 9 <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/FT.html#FTGRA>.
- 10 <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/PH.html#PHFAX>.
- 11 For a graphical representation of the model see Scholger 2019, 46.
- 12 Art & Architecture Thesaurus (AAT), accessed February 8, 2019, <http://www.getty.edu/research/tools/vocabularies/aat/>.
- 13 Artwork list of Hartmut Skerbisch, accessed February 8, 2019, <http://gams.uni-graz.at/o:sker.artworks>.
- 14 Taxonomy on concepts mentioned by the artist, accessed February 8, 2019, <http://gams.uni-graz.at/o:sker.concepts>.
- 15 TEI-Facsimile-Plugin, accessed February 8, 2019, <https://github.com/oxygenxml/TEI-Facsimile-Plugin>.

- 16 TEI-Zoner, accessed February 8, 2019, <http://teicat.huma-num.fr/zoner.php>.
- 17 <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/ref-path.html>.
- 18 <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/ref-label.html>.
- 19 <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/ref-material.html>.
- 20 <http://www.tei-c.org/Vault/P5/3.5.0/doc/tei-p5-doc/en/html/ref-att.global.analytic.html>.
- 21 IIIF Image API, accessed February 8, 2019, <http://iiif.io/api/image/2.0/>.
- 22 Alexander Fleming's Lab Books, British Library, Add MS 56106–56225, accessed May 25, 2018, <https://www.bl.uk/collection-items/alexander-flemings-lab-books>.
- 23 Sketches and notes by Albrecht Dürer, British Library, Add MS 5231, accessed May 25, 2018, http://www.bl.uk/manuscripts/FullDisplay.aspx?ref=Add_MS_5231.
- 24 edition humboldt digital, accessed May 25, 2018, <https://edition-humboldt.de/>.

AUTHOR

MARTINA SCHOLGER

Martina Scholger is a senior scientist and researcher at the Centre for Information Modelling – Austrian Centre for Digital Humanities at the University of Graz. She studied art history and received her PhD in digital humanities with a study on artists' creation processes documented in their notebooks. She is teaching information and text modelling and is involved in numerous cooperation projects in the field of digital scholarly editing and text mining. She has been a member of the Institute for Documentology and Scholarly Editing since 2014 and a member of the TEI Technical Council since 2016, where she is currently serving as Chair.