



Journal of the Text Encoding Initiative

Issue 9 | September 2016 - December 2017
Selected Papers from the 2014 TEI Conference

Towards a Model for Encoding Correspondence in the TEI: Developing and Implementing `<correspDesc>`

Peter Stadler, Marcel Illetschko and Sabine Seifert



Electronic version

URL: <http://journals.openedition.org/jtei/1433>

DOI: 10.4000/jtei.1433

ISSN: 2162-5603

Publisher

TEI Consortium

Electronic reference

Peter Stadler, Marcel Illetschko and Sabine Seifert, « Towards a Model for Encoding Correspondence in the TEI: Developing and Implementing `<correspDesc>` », *Journal of the Text Encoding Initiative* [Online], Issue 9 | September 2016 - December 2017, Online since 24 September 2016, connection on 19 April 2019. URL : <http://journals.openedition.org/jtei/1433> ; DOI : 10.4000/jtei.1433

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

Towards a Model for Encoding Correspondence in the TEI: Developing and Implementing <correspDesc>

Peter Stadler, Marcel Illetschko, and Sabine Seifert

1. Current Digital Editions of Correspondence Material

1.1 Digital Editions and Main Questions

- ¹ For several years now, digital projects and editions have been dealing with many varieties of correspondence. The encoding of letters ¹ and the methodological reflection on this process have a long tradition in the TEI. One early achievement in this area was the *Model Editions Partnership: Historical Editions in the Digital Age*² (MEP), developed during the 1990s, which provided several sample editions of historical documents, including letters. On the basis of these, the project editors reflected on new approaches and a general model for preparing digital editions. Working with SGML (later XML) and TEI, a number of custom MEP elements were added to the official TEI elements for the marking up of letter-specific phenomena, for example, the sender and addressee, typical closing phrases, or a postscript (<mep:sender>, <mep:addressee>, <mep:closing>, <mep:ps>).³

- 2 Current outstanding correspondence projects include *DALF: Digital Archive of Letters in Flanders* (DALF)⁴ and *Vincent Van Gogh: The Letters*,⁵ the latter of which is, in many respects, an admirable example of a digital edition. In addition to the emergence of these significant correspondence projects, efforts to conceptualize and standardize the encoding of correspondence within the TEI itself began to become more visible. For example, a feature request for a “letters/memos module” was submitted in 2004. This request generated interest, but was not realized.⁶
- 3 Numerous digital editions of letters are currently in development. The *Alfred Escher Letters Edition*⁷ launched with extensive revisions in July 2015, and *BurckhardtSource*⁸ transitioned from a beta version to an official release in September 2015. Other active projects include *Carl Maria von Weber –Collected Works (WeGA)*,⁹ *Letters and Texts: Intellectual Berlin around 1800*,¹⁰ *Letters of 1916*,¹¹ *The Walt Whitman Archive*,¹² *The William F. Cody Archive*,¹³ and *August Wilhelm Schlegel’s Correspondence*.¹⁴ All of these projects struggled with the specific requirements of encoding correspondence, and most developed new elements or borrowed some from other projects for situations not adequately covered by the TEI Guidelines.
- 4 Projects currently in development but not yet online include, for example, *The Complete Letters of Willa Cather*¹⁵ and *Epistolary Networks*.¹⁶ Besides editions, interfaces for editorial work or correspondence-specific search queries have been developed, such as *correspSearch*¹⁷ or *Early Modern Letters Online* (EMLO).¹⁸ All of these editions, interface projects, and also visualization tools such as *Mapping the Republic of Letters*¹⁹ or *Visual Correspondence*,²⁰ rely heavily on capturing precisely the metadata of each letter.
- 5 This short overview of editions and projects is by no means exhaustive.²¹ In all of these endeavors, much technical development has been done and many elaborate encodings have been produced. However, the lack of common, community-endorsed encoding guidelines, especially for letters and other types of correspondence, has resulted in much duplicated work, often hugely varying encoding practices, and obstacles—caused by different encoding schemes—to the interchange of texts. This situation leads to two main questions about digital editions of correspondence:
- What changes to the current TEI Guidelines are needed to better support (and standardize) the encoding of correspondence?

- And, going a step further, how can correspondence editions most effectively be *linked* to one another?

1.2 Two Inspiring Examples: DALF and WeGA

- 6 When DALF and WeGA each started, the TEI Guidelines were extensive and covered a wide variety of ways to encode manuscripts, but specific encoding guidelines for epistolary material were still missing. Because of this, both projects tried to address the problem by developing their own customizations and documentation for encoding metadata and text transcriptions. The Belgian DALF project,²² a pioneer of encoding correspondence, was launched by the Centre for Scholarly Editing and Document Studies (CTB). In 2003, the *DALF Guidelines for the Description and Encoding of Modern Correspondence Material* (Vanhoutte and Van den Branden 2003) were published, with the encoding based on TEI P4. The DALF customization was extensively documented and served as a model for several correspondence projects thereafter. In 2012, a trial update to TEI P5 was tested in conjunction with a re-evaluation of the DALF guidelines. They followed two principles: first, changing existing TEI P5 elements as little as possible, and second, isolating the project-specific changes from the official TEI P5 elements whenever possible. This new customization was documented in “DALF: A Preliminary P5 Proposal” (Van den Branden 2013).
- 7 In 2011, before DALF’s trial update to P5, the *Carl Maria von Weber—Collected Works* (WeGA) project, founded at the University of Paderborn and funded by the Academy of Sciences and Literature Mainz, had launched with access to many manuscripts and letters. TEI P5 was used right from the beginning and in the same year, “TEI P5 for Correspondence: A Recommendation for the Encoding of Correspondence Material” (Stadler 2011) was developed and made available to the community. Subsequently, all development of the WeGA correspondence schema was publicly accessible at GitHub²³ and became a model for several correspondence projects that used TEI P5 from the outset.
- 8 In keeping with best practice, both projects tried to keep the changes to the existing TEI elements as minimal as possible. They defined a special wrapper element, firstly in order to keep correspondence metadata in a single place for convenient encoding and querying, and secondly to keep the rest of the TEI header untainted by these additions. In each case, the correspondence

description was understood as forming a part of the description of the source (that is a manuscript, print document, or digital file). Hence the wrapper element became part of the <sourceDesc> element.

1.2.1 Encoding Example of *Digital Archive of Letters in Flanders*

9 Comparing the correspondence wrapper elements in both DALF and WeGA shows that the structure and the introduced child elements differ slightly. In the DALF project, the wrapper element <dalf:letDesc> (letter description) “[g]roups together all letter-specific metadata for a DALF document.”²⁴ The main child element is <dalf:letHeading> which “[c]ontains a structured description of bibliographical information of a letter.”²⁵ These consist of four key metadata fields that are defined as follows, each with new elements:

- <dalf:letAuthor>: the author of the letter (no differentiation was made between the notions of an “author” and a “sender”),
- <dalf:letAddressee>: the addressee of the letter,
- <dalf:letPlace>: the place where the letter was written, and
- <dalf:letDate>: the date of the letter’s origin.

10 Information on parties other than the author who were responsible for the content of the letter are encoded within the already existing TEI element <respStmt> within <dalf:letHeading>.

11 Besides these basic characteristics of a letter in <dalf:letHeading>, additional information can be provided in other new elements. These are restricted to:

- <dalf:type>: the formal classification of the letter,
- <dalf:env0cc>: the occurrence of envelopes,
- <dalf:fig0cc>: the occurrence of illustrations in the letter.

12 Extra notes on the letter can be added in the TEI element <note>. All other information, for example on provenance, physical appearance, or history of the letter, are put in <msDesc> (manuscript description), as there are already sufficient TEI elements with which to encode such data.

- 13 Within <sourceDesc> the grouping is thus first <biblStruct>, second <dalf:letDesc>, and third <msDesc>. [Example 1](#) shows the encoding for a letter from the DALF corpus (with a repeated <dalf:letHeading>, since there are actually two letters on one sheet of paper in this case):

Example 1. Encoding example from DALF (Van den Branden 2013, file `dalfP5test.xml`).

```
<dalf:letDesc xmlns:dalf="http://ctb.kantl.be/DALF/2.0">
  <!-- 5.1 Composite Letters: multiple letHeading elements for each subdivision
in the letter -->
  <dalf:letHeading xml:id="KVDW.part" default="true">
    <dalf:letAuthor ref="names.xml#KVDW" attested="yes">Karel van de Woestijne</
dalf:letAuthor>
    <dalf:letAddressee ref="names.xml#EDB" attested="yes">Emmanuel de Bom</
dalf:letAddressee>
    <dalf:letPlace attested="no"/>
    <dalf:letDate notAfter="1904-07-25" notBefore="1904-08-02">na 1904-07-25 en
voor 1904-08-02</dalf:letDate>
    <!-- 5.1 Composite Letters: multiple letHeading elements for each subdivision
in the letter -->
  </dalf:letHeading>
  <dalf:letHeading xml:id="MVDW.part">
    <dalf:letAuthor ref="names.xml#MVDW" attested="yes">Mariette van de
Woestijne</dalf:letAuthor>
    <dalf:letAddressee ref="names.xml#EDB" attested="yes">Emmanuel de Bom</
dalf:letAddressee>
    <dalf:letPlace attested="no"/>
    <dalf:letDate notAfter="1904-07-25" notBefore="1904-08-02">na 1904-07-25 en
voor 1904-08-02</dalf:letDate>
  </dalf:letHeading>
  <dalf:env0cc occ="false"/>
  <dalf:fig0cc quantity="1"/>
  <dalf:type>Brief, geen omslag</dalf:type>
</dalf:letDesc>
```

1.2.2 Encoding Example of *Carl Maria von Weber—Collected Works*

- 14 The WeGA offers a different, shallower structure which groups the key metadata parallel to each other directly within the wrapper element called <wega:correspDesc> (correspondence description). The <wega:correspDesc> wrapper “groups together meta data about the (historic) correspondence such as sender, addressee etc.” (Stadler 2011). The variety of the selected metadata also differs slightly. Newly defined elements are:
- the sender, encoded within the element <wega:sender> (here, a differentiation was made between an “author” and a “sender”),
 - the addressee in the element <wega:addressee>,
 - the place where the letter was written in <wega:placeSender>,
 - the date when the letter was written in <wega:dateSender>,
 - the place where the letter was received in <wega:placeAddressee>,
 - the date when the letter was received in <wega:dateAddressee>, and
 - the position of the letter within the whole thread of correspondence in <wega:context>.
- 15 For capturing the beginning of the letter, the already existing element <incipit> was borrowed from the manuscript description and included within <wega:correspDesc>. As in the DALF guidelines, there is no encoding of information on the source and its history, which remain in the manuscript description.
- 16 Within the source description, the first element is <wega:correspDesc>, and the second contains the bibliographic information (<msDesc>, <listWit>, <biblStruct>, or similar). Example 2 shows the encoding for a letter from WeGA:

Example 2. Encoding example from WeGA (file A041001.xml from Stadler 2011).

```
<wega:correspDesc xmlns:wega="http://www.weber-gesamtausgabe.de">
  <wega:sender>
    <persName key="A002068">Weber, Carl Maria von</persName>
  </wega:sender>
  <wega:addressee>
    <persName key="A000213">Weber, Caroline</persName>
  </wega:addressee>
  <wega:placeSender>
    <placeName>Berlin</placeName>
  </wega:placeSender>
  <wega:placeAddressee>
    <placeName>Prag</placeName>
  </wega:placeAddressee>
  <wega:dateSender>
    <date from="1817-01-01" to="1817-01-04" n="01">1. bis 4. Januar 1817</date>
  </wega:dateSender>
  <incipit>Prost Neujahr!! Millionen Bußen. und gute Nacht</incipit>
  <wega:context>
    <ref target="http://www.weber-gesamtausgabe.de/A041003" n="next">Brief Webers
an Caroline Brandt vom 5.–7. Januar 1817</ref>
  </wega:context>
</wega:correspDesc>
```

- 17 It goes without saying that there are several possible ways to encode the many different types of correspondence (letters, postcards, and so on), and each approach needs to fit the respective project’s aims. DALF and WeGA present detailed and exemplary models. These two encoding models and editions were the basis of the work of the TEI Correspondence Special Interest Group presented below.

1.3 The TEI Correspondence SIG’s Task Force “correspDesc”

- 18 In 2013, the TEI Correspondence SIG put together a task force called “correspDesc” to address the question of how to modify the TEI in order to better support the encoding of correspondence. This task force consisted of the three authors of this article, who have been co-conveners of the Correspondence SIG since 2014. Our goal was to create a formal proposal for a correspondence

metadata element (with appropriate child elements) for approval by the TEI Council as a new feature of the TEI standard. Dealing with transcriptional elements was set aside as a second step. We put the focus on standardization of the editorial information and its encoding within the <teiHeader>, and discussed and selected the “core” correspondence-specific metadata that should be encoded—and thus be conveniently identifiable—in one fixed location in the TEI header.

- 19 The development of this model for encoding correspondence-specific metadata was inspired largely by the work of DALF and WeGA, for these projects had offered the most successful and “reusable” customizations thus far. We built on these customizations rather than starting from scratch and began with an evaluation of the overlaps and differences between the two customizations. We followed both DALF and WeGA in their two basic assumptions: on the one hand, the TEI standard should be modified cautiously by adding only a few elements while not changing existing elements. On the other hand, there should be a new wrapper element to store the key metadata on the encoded piece of correspondence instead of scattering this information throughout the TEI header. This wrapper element is called <correspDesc> (correspondence description), and it covers correspondence-specific metadata only. All information describing the manuscript (or any other text-bearing object) still resides within the manuscript description <msDesc>. This also means that the element <correspDesc> alone does not provide a complete description of a letter or other piece of correspondence, but gives *selected* correspondence-specific data. A full description of a given letter is provided by <correspDesc> in conjunction with the manuscript description in <msDesc>.
- 20 During the course of developing <correspDesc> and refining the underlying definition of correspondence in general, the resulting correspondence-specific guidelines and the setup of the added TEI elements moved somewhat away from the concepts of DALF and WeGA. This, however, does not diminish their importance for conceptualizing <correspDesc> and the encoding guidelines for correspondence.

2. Theory

- 21 As outlined above, providing better support (and standardization) for the encoding of correspondence was one of the central goals of the task force. The process of conceptualizing was guided by two things: first, a theory of correspondence in general, and second, the possible application of existing TEI elements (as well as the need for new elements) for a given set of letters or other pieces of correspondence.

2.1 Theory of Correspondence in General

- 22 We first had to settle on what constitutes a letter or any other piece of correspondence, e.g., a postcard or a telegram. Inspiration came from many different sources, the significant ones being Halsband (1958), Bluhm and Meier (1993), Roloff (1998), Barton and Hall (2000), Zeller (2002), How (2003), Stanley (2004), Wiethölter and Bohnenkamp (2010), Berg (2011), Bohnenkamp and Richter (2013), and Hankins (2015). Three aspects emerged as the central ones: The materiality of a piece of correspondence, the text of the message, and the “eventness” of the communicative act.

2.1.1 The Letter as an Object

- 23 It seems obvious that any piece of correspondence is tightly bound to its materiality. First, the available space on a text-bearing object not only limits the amount of the text, it also puts stylistic constraints on the message; compare, for example, a postcard with a letter, where on the former you will not usually find such elaborate openers or closers as on the latter. Second, there are implications for the communicative act, since the object is part of the (social contexts of the) message—was expensive paper used or just some scrap of paper?—and some material forms of correspondence conceal their contents while others bear text visibly.
- 24 Since these material aspects play a distinct role for many editions of texts in general (editions of correspondence being a subset of these), the TEI fortunately already features a full-fledged module for the encoding of manuscripts (and other text-bearing objects in general) with <msDesc> (cf. Pierazzo 2011). The particular characteristics of correspondences—those material aspects that are genuinely related to correspondence (e.g., attachments, enclosures, and envelopes)—can also be encoded within <sourceDesc>. A taxonomy of media types (e.g., postcard, letter, or email) may be specified in the <profileDesc>.

2.1.2 The Text of the Letter

- 25 For the encoding of textual content of a message, the TEI Guidelines (TEI Consortium 2015) provide a most comprehensive tag set. One can elaborately encode names, dates, places, and their relationships. One can, as well, document all sorts of editorial interventions and features of the copy text. Again, particular textual characteristics of correspondence can already be dealt with. For example, the existing elements <postscript>, <opener>, and <closer> provide for encoding the prototypic text structure of a letter. Of course, this prototypic structure does not apply to all letters, nor to correspondence in general, so the aforementioned elements and their applicability are constantly called into question. Yet, as pointed out in section 1.3, this issue was not part of the task force's mandate but needs to be addressed in a subsequent step.

2.1.3 The Letter as an Event

- 26 Besides the material and the textual features of a message, its “eventness” is of exceptional importance. In general, an event introduces change to the associated parties and to the particular communication continuum in which the correspondence takes place (Bohnenkamp and Richter 2013, 4). Each individual message is not a sole entity but a reaction to some previous message, and triggers a reaction itself. For some, this “temporal sequence of the Before and After with something happening in the middle of those” plays a predominant role: “More important than the question of what happens is the fact that something happens, that is the mere carrying out”²⁶ (Stenger 2010, 30). In fact, the mere (written) text of a message was not emphasized until the seventeenth and eighteenth centuries, Stenger argues. Very often, “the surrounding circumstances—sent works, presents, works of art, other oral messages—show that the letter is embedded within a whole ensemble of communication media that made correspondence a multimedia process”²⁷ (ibid., 32).
- 27 Correspondence is frequently called a “half dialogue,” a “conversation amongst absentees”; Janet Gurkin Altman, for example, speaks about “the letter’s function as a connector between two distant points, as a bridge between sender and receiver” (Altman 1982, 13). This view has much in common with the classic model of communication as laid out by Claude Elwood Shannon and Warren Weaver in the 1940s (Shannon and Weaver 1949). The so-called Shannon–Weaver model can be reduced to the entities sender, receiver, and a connecting transmission.

28 We therefore tried to incorporate the basic principles of this communication model into our concept of what constitutes correspondence. Although, of course, not every letter follows the same basic pattern, it nevertheless seemed reasonable to concentrate on these key points:

- sender
- receiver
- origin location
- destination location
- message(s) before and after
- date sent
- method of transmission
- date received

29 Hence, the proposed correspondence description needed to provide information about persons (or organizations) as sender, receiver, or messenger. In addition to this, it needed to support the encoding of the respective dates and places as well as to provide a mechanism that would point at (or reference) preceding and subsequent messages.

2.2 Theory of Correspondence and TEI Entities

30 The previously outlined “eventness” of correspondence, that is, the communicative act associated with correspondence, can be decomposed into various “actions.” These actions (which include sending, receiving, and transmitting) form the atomic events of a communicative act and are associated with people, dates, and places.

2.2.1 Persons

31 An act of communication has a sender of a message who is not necessarily identical to the author of the text of the message. In the TEI Guidelines, “<author> in a bibliographic reference, contains the name(s) of an author, personal or corporate, of a work; for example in the same form as that provided by a recognized bibliographic name authority.” (TEI Consortium 2015, “<author>”)²⁸ Obviously, this definition is connected to specific concepts of “authorship” and “work” which cannot be discussed in great detail here. Of course, it is possible in an act of communication or

correspondence to send a copy of a Shakespearean poem instead of writing an original love letter; it is possible for a writer to send a publishing contract in order to discuss it with a friend; and it is possible to send a disembodied ear in order to threaten an enemy instead of writing a message at all. Still, correspondence and communication happen in each of these instances. These sorts of possibilities make it crucial to differentiate between the “author” and the “sender” of a letter. If author and sender are the same person—which in most letters will be the case—one can use a pointer to the corresponding part of the TEI header to indicate that.

- 32 Furthermore, correspondence generally has one or more intended (perhaps fictional) “addressees” or “receivers.” It is important to point out that the addressee “spoken” to by the piece of correspondence is not necessarily the same person as the one who actually “receives” the letter. One might also want to encode persons involved in the transmission of the letter, such as a dear friend entrusted with the letter, a personal messenger, or an official letter carrier.

2.2.2 Dates

- 33 A piece of correspondence generally has certain points in time when it was written and sent which are not necessarily the same, and may even be unknown, approximate, or indeterminate. One could also argue that the date of receipt of a letter is documented in some cases and should thus be encoded as well.

2.2.3 Places

- 34 Similarly, a piece of correspondence is generally sent *from* a particular location, which may or may not be the same as the place(s) of composition. A letter is also typically sent *to* a specific location. Very often these bits of information are evident in the text of the message, in an address line, postmark, or in the automatically generated header metadata in an email message, so they can and should be captured in the encoding.

2.2.4 Transmitting, Redirecting, Forwarding

- 35 A piece of correspondence generally has a process or a medium and/or one or more executor(s) of transfer (messenger, carrier, postman, carriage, fax machine, Internet) and may be redirected (usually unread/without acknowledgement of the content) or forwarded (usually read/with acknowledgement of the content). Again, this vital information should be recorded in the encoding.

2.2.5 Context

- 36 Typically, a piece of correspondence is not an isolated entity but a (written) act of communication within a communication continuum (see Bohnenkamp and Richter 2013, 4), in which the correspondence is defined by its relative position between messages sent “before” and “after.” Frequently, a piece of correspondence is sent as a reaction to another piece of correspondence and triggers an answer itself. This communicative thread does not necessarily have a simple chronological order, but may contain overlaps due, for example, to postal delivery issues or procrastination. Establishing this succession is a common editorial task; thus, the envisioned encoding model should accommodate detailed description of context.

3. Putting Theory into Practice

- 37 Starting from the main question, “What changes to the current TEI Guidelines would be needed to support scholarly encoding of correspondence?” and after considering the above-mentioned communicational aspects, we finally tried to define as few new correspondence-related elements as necessary while maximizing reliance on the already existing TEI framework. In spite of our initial idea of just implementing dedicated elements such as <sender>, <addressee>, <placeSender>, or <transmission> (based on the customizations used by DALF and WeGA), we were influenced by the arguments of some contributors on the TEI Correspondence SIG list. These contributors suggested that it would be more appropriate to understand correspondence as an involvement of various persons and responsibilities rather than just sender, addressee, or transmitter, and, therefore, that it would be best to use one overall term (<participant>) with different roles (e.g., @role = “author”, “sender”, “signer”, or “co-signer”). We created encoding examples with this model but soon reached its limitations because of the sole emphasis on persons. However, in combination with a different suggestion—namely to wrap all information about the sending (or receiving) side of the correspondence in one element each and to encode names, dates, and places using existing TEI elements—we finally managed to develop a specification that seemed both theoretically justifiable and practically useful.

3.1 <correspDesc>

- 38 The direct implementation of our communication-oriented concept of correspondence starts with the new element <correspDesc> (correspondence description) that stores the key metadata about the communicative act. This element has one or more <correspAction> and <correspContext> elements as children. All other contents are supplied using already existing elements.
- 39 Initially, we wanted to include <correspDesc> in <sourceDesc> (and by that means make it model.biblLike in order to create, for example, <listBibl>s for collections of letters). After discussions with the TEI Council, we finally decided that <profileDesc> would be more appropriate. In spite of the “bibliographical” character of information like sender, date, and place of writing, correspondence-related metadata puts the emphasis not on specifying the source but on describing the text and its several aspects. It is therefore better described by the “non-bibliographic aspects of a text, ... the situation in which it was produced, the participants and their setting” (TEI Consortium 2015, “<profileDesc>”).²⁹
- 40 Below is a very simple example of an encoded letter using the new <correspDesc> which “contains a description of the actions related to one act of correspondence” (TEI Consortium 2015, “<correspDesc>”).³⁰

Example 3. Encoding example of the TEI element <correspDesc> (TEI Consortium 2015, “<correspDesc>”).

```
<correspDesc>
  <correspAction type="sent">
    <persName>Carl Maria von Weber</persName>
    <settlement>Dresden</settlement>
    <date when="1817-06-23">23 June 1817</date>
  </correspAction>
  <correspAction type="received">
    <persName>Caroline Brandt</persName>
    <settlement>Prag</settlement>
  </correspAction>
  <correspContext>
    <ref type="prev" target="http://www.weber-gesamtausgabe.de/A041209">Previous
letter of <persName>Carl Maria von Weber</persName> to <persName>Caroline Brandt</
persName>: <date from="1817-06-19" to="1817-06-20">June 19/20, 1817</date>
    </ref>
    <ref type="next" target="http://www.weber-gesamtausgabe.de/A041217">Next
letter of <persName>Carl Maria von Weber</persName> to <persName>Caroline Brandt</
persName>: <date when="1817-06-27">June 27, 1817</date>
    </ref>
  </correspContext>
</correspDesc>
```

3.2 <correspAction>

41 <correspAction> is the “heart” of the new model for encoding correspondence with the TEI. It is the “atomic unit” relating to events of a given communicative act and gives flexible opportunities to include all the information about associated people, dates, and places as described in section 2 by using existing TEI elements. The <correspAction> element “contains a structured description of the place, the name of a person/organization and the date related to the sending/receiving of a message or any other action related to the correspondence.” (TEI Consortium 2015, “<correspAction>”)³¹ Suggested values for its @type attribute are:

- "sent" (information concerning the sending or dispatch of a message),
- "received" (information concerning the receipt of a message),

- "transmitted" (information concerning the transmission of a message, that is, between the dispatch and the next receipt, redirect, or forwarding),
- "redirected" (information concerning the redirection of an unread message), and
- "forwarded" (information concerning the forwarding of a message).

42 Example 3 above shows how to encode a single letter with the actions "sent" (a known author/writer/sender sends the letter from a known place on a known date) and "received" (a known addressee/receiver receives the letter at a known place—but on an unknown date). Information about the context in which the letter was sent is also supplied (in this case by referencing the previous and next letters of the author/sender).

3.3 <correspContext>

- 43 The element "<correspContext> (correspondence context) provides references to preceding or following correspondence related to this piece of correspondence" (TEI Consortium 2015, "<correspContext>").³² It therefore identifies the proper place of a particular piece of correspondence (a letter, for example) in the communication continuum, as it is defined by its relative position between messages sent "before" and "after." This may be very useful in capturing the correspondence network of a single person, where different letters written on the same day may be part of different discussions with different addressees, or where one and the same (forwarded) letter generates different answers from different writers.
- 44 Sometimes one will have to deal with pieces of correspondence within other texts (such as biographies) and use a combination of <div> and <correspDesc>. Very often it will be necessary to combine many different letters, for instance, while encoding an anthology. One could then encode the text of each letter in a single <div> in the <body> and link to the corresponding <correspDesc> (one for each letter) in the <teiHeader>. For facilitating this use case, <correspDesc> is made a "declarable element" (TEI Consortium 2015, sec. 15.3.2, "Declarable Elements"),³³ and thus can be linked to by means of a @decls attribute on the corresponding element within <text>, as demonstrated in the following example:

Example 4. Encoding example of an anthology of letters (TEI Correspondence SIG GitHub repository,

<correspDesc> example 7)³⁴

```

<teiHeader>
  <profileDesc>
    <!-- Every letter gets its own correspDesc -->
    <correspDesc xml:id="letter1">
      <correspAction type="sent">
        <persName>Amalia Beer</persName>
        <placeName>Berlin</placeName>
        <date when="1815-02-28">28 February 1815</date>
      </correspAction>
      <correspAction type="received">
        <persName>Giacomo Meyerbeer</persName>
        <placeName>Paris</placeName>
      </correspAction>
    </correspDesc>
    <correspDesc xml:id="letter2">
      <correspAction type="sent">
        <persName>L. Schoenberger</persName>
        <placeName>London</placeName>
        <date when="1815-03-04">4 March 1815</date>
      </correspAction>
      <correspAction type="received">
        <persName>Giacomo Meyerbeer</persName>
        <placeName>Paris</placeName>
      </correspAction>
    </correspDesc>
  </profileDesc>
</teiHeader>
<text>
  <body>
    <!-- We put every letter in a <div> of its own and link to the respective
<correspDesc> via @decls -->
    <div decls="#letter1">
      <opener>
        <dateline>Berlin den 28t Februar <choice><sic>1814</sic><corr>1815</corr></
choice></dateline>
      </opener>

```

```

    <p>Mein lieber Sohn, ...</p>
  </div>
  <div decls="#letter2">
    <opener>
      <dateline>London d. 4. März <supplied>1</supplied>815</dateline>
      <salute>Schätzbarster Freund!</salute>
    </opener>
    <p>Wir sind ganz wohlbehalten ...</p>
  </div>
</body>
</text>

```

- 45 More examples of suggested encodings as well as unusual encoding challenges (such as emails, multiple senders/receivers, fictional letters, redirections, a single document comprising two acts of correspondence, or one act of correspondence with multiple witnesses) can be found at [the SIG's GitHub repository](#).³⁵

4. Interchange Format

- 46 A crucial aspect of the development of <correspDesc> has been the wish to facilitate interchange—even interoperability—of the metadata from encoded correspondence texts, as there is a general and growing demand from correspondence projects for interchange and linked-data capabilities. To provide for this, we developed a model of the <correspDesc> element in a concentrated form that is essentially a constrained subset of the full TEI standard. This Correspondence Metadata Interchange (CMI) format relies heavily on authority files and external standard formats. Authority files like the [Integrated Authority File \(GND\)](#)³⁶ or the [Virtual International Authority File \(VIAF\)](#)³⁷ are used for identifying persons and places, and the stricter [W3C format](#)³⁸ (being a subset of ISO 8601) is used for the encoding of dates. Including such standards can introduce new problems, however, when there are several authority files for one piece of information or when there are none.
- 47 Nevertheless, along with the development of the CMI format, a corresponding Web service using this customization was created by Stefan Dumont at the Berlin-Brandenburg Academy of Sciences and Humanities. This Web service called [correspSearch](#)³⁹ makes the correspondence-

specific metadata of (to this day) seven German- and French-language correspondence editions searchable with one query. This service gives an idea of what is possible when correspondence projects are linked. The CMI format and the *correspSearch* Web service, still under development, aim to address in greater detail the question of linking and interchange between correspondence projects.

5. Organizational Difficulties and Achieving Official TEI Status

- 48 As already noted in [section 1](#), a TEI letters/memos module was previously requested in 2004. It was deferred in 2007 (with the advent of TEI P5) with a note by Syd Bauman, a member of the TEI Technical Council: “I still think it is quite a good idea to have better support for letters & memos in TEI, and I am hopeful that the original poster, the DALF folks, and others will help make this a reality for a subsequent release of P5.”⁴⁰ There is almost no information documented on the TEI Sourceforge ticket tracker for the time between the creation of the ticket and the closing of the ticket in 2007, so most probably the initial momentum was lost and not much activity followed.
- 49 In 2008, the Correspondence SIG was established and held its first meeting during the TEI members’ meeting in London. That meeting was rather successful in reaching out to and involving various people and projects. However, it took another year before the first steps were taken. At the 2009 TEI meeting in Ann Arbor, Michigan, an initial task force was set up consisting of Markus Flatscher, Bert Van Raemdonck, and Peter Stadler. Their goal was to map the DALF TEI P4 customization to TEI P5 as a basis for further work on a correspondence customization; hence the code name “Dalfy.” One of the problems for this task force was the transatlantic makeup of the task force, with Markus Flatscher located in Virginia (USA) and the others located respectively in Belgium and Germany. Modern telecommunication offers a variety of tools for conference calls, but different time zones limit the intersection of work hours, and face-to-face meetings remain a much more productive way to embark on such a venture. Some efforts were made to acquire funding, but the task force failed to secure funding and no “Dalfy” mapping was developed.

- 50 At the Correspondence SIG meeting during the 2012 TEI Conference in College Station, Texas, a spontaneous “hack session” led to the “[second draft for a correspondence ODD](#).”⁴¹ At the 2013 TEI members’ meeting in Rome, the second task force named “correspDesc” was established consisting of—as mentioned above—the three authors of this paper. The organizational preconditions were more favorable this time since all members were native German speakers and several face-to-face meetings could be arranged in Berlin thanks to the generous funding of home institutions. This time, the goal was explicitly to continue the work that had been started in College Station and to create a formal proposal for a <correspDesc> element for approval by the TEI Council as a new element of the TEI standard. A GitHub repository for developing the ODD customization and providing access to examples and documentation⁴² was set up, and the task force’s work was documented on the SIG’s wiki space.
- 51 The task force sought but did not receive substantial feedback from the wider TEI community during our work on the <correspDesc> proposal. This may have had to do with the communication channels we chose to use for interacting with the wider community: besides the SIG’s wiki page and the GitHub repository, we used the SIG and the TEI mailing lists and were co-hosts of the first ever TEI tweet chat. Perhaps a series of workshops would have helped us acquire additional input from other domain experts.⁴³ The feedback we *did* receive during the process was quite encouraging, though, and in June 2014 we had the proposal—that is, the formal specification with documentary prose and examples—in good enough shape to open a [feature request](#) on the TEI Sourceforge ticket tracker.⁴⁴ This initiated the official process of integrating <correspDesc> and related elements into the *TEI Guidelines*. Fortunately, one member of our task force was also a member of the TEI Council during this period, so the Council was constantly reminded about this issue, and queries about the proposal could be answered instantly by the task force member also serving on the TEI Council. Nonetheless, the TEI Council had some concerns and suggestions, and the proposal required a few revisions to address the Council’s concerns and obtain approval. In November 2014, this issue was finally discussed during the Council’s face-to-face meeting at Duke University and accepted for implementation into the TEI Guidelines. It was duly implemented in February 2015 and found its way into the official TEI standard with the TEI P5 release 2.8.0 in April 2015.

52 We would like to thank all members of the greater TEI community and the Correspondence SIG who took part in the development of the correspDesc-proposal with feedback and advice. Special thanks are due to the TEI Council of the term 2014–15 and to Ron Van den Branden for continuous support and counseling.

BIBLIOGRAPHY

- Altman, Janet Gurkin. 1982. *Epistolarity: Approaches to a Form*. Columbus: Ohio State University Press.
- Barton, David, and Nigel Hall, eds. 2000. *Letter Writing as a Social Practice*. Amsterdam: John Benjamins Publishing Co.
- Berg, Temma. 2011. "Truly Yours: Arranging a Letter Collection." *Eighteenth-Century Life* 35(1): 29–50. doi:10.1215/00982601-2010-026.
- Bluhm, Lothar, and Andreas Meier, eds. 1993. *Der Brief in Klassik und Romantik: Aktuelle Probleme der Briefedition*. Würzburg: Königshausen & Neumann.
- Bohnenkamp, Anne, and Elke Richter, eds. 2013. *Brief-Edition im digitalen Zeitalter*. Berlin: De Gruyter.
- Halsband, Robert. 1958. "Editing the Letters of Letter-Writers." *Studies in Bibliography* 11: 25–37.
- Hankins, Gabriel. 2015. "Correspondence: Theory, Practice, and Horizons." In *Literary Studies in the Digital Age*, edited by Kenneth M. Price and Ray Siemens. Accessed March 11, 2016. <https://dlsanthology.commons.mla.org/correspondence-theory-practice-and-horizons/>.
- How, James. 2003. *Epistolary Spaces: English Letter-writing from the Foundation of the Post Office to Richardson's Clarissa*. London: Ashgate.
- Pierazzo, Elena. 2011. "A Rationale of Digital Documentary Editions." *Literary and Linguistic Computing* 26 (4): 463–77. doi:10.1093/lc/fqr033.
- Roloff, Hans-Gert, ed. 1998. *Wissenschaftliche Briefeditionen und ihre Probleme. Editionswissenschaftliches Symposium*. Berlin: Weidler.
- Stadler, Peter. 2011. "TEI P5 for Correspondence: A Recommendation for the Encoding of Correspondence Material." Accessed June 30, 2015. <http://wiki.tei-c.org/index.php/File:CorrespondenceProposal.zip>, available at http://wiki.tei-c.org/index.php/SIG:Correspondence/ODD_work.
- Stanley, Liz. 2004. "The Epistolarium: On Theorizing Letters and Correspondences." *Auto/Biography* 12 (3): 201–35. doi:10.1191/0967550704ab0140a.
- Stenger, Jan. 2010. "Reich an Worten, arm an Inhalt? Der spätantike Brief als Ereignis." In *Der Brief—Ereignis & Objekt*, edited by Waltraud Wiethölter and Anne Bohnenkamp, 26–41. Frankfurt am Main: Stroemfeld.

- TEI Consortium. 2015. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 2.9.1. Last updated on October 15. <http://www.tei-c.org/Vault/P5/2.9.1/doc/tei-p5-doc/en/html/>.
- Vanhoutte, Edward, and Ron Van den Branden, eds. 2003. *DALF Guidelines for the Description and Encoding of Modern Correspondence Material*, version 1.0. Gent: Centrum voor Teksteditie en Bronnenstudie, Koninklijke Academie voor Nederlandse Taal- en Letterkunde. Revised May 4, 2005. <http://www.kantl.be/ctb/project/dalf/dalfdoc/>.
- Shannon, Claude Elwood, and Warren Weaver. 1949. *The Mathematical Theory of Communication*. Urbana: Univ. of Illinois Press.
- Van den Branden, Ron. 2013. "DALF: A Preliminary P5 Proposal." August 21. http://ctb.kantl.be/project/dalf/P5/DALF_P5-p0.1.zip.
- Wiethölter, Waltraud, and Anne Bohnenkamp, eds. 2010. *Der Brief—Ereignis & Objekt: Frankfurter Tagung*. Frankfurt am Main: Stroemfeld.
- Zeller, Hans. 2002. "Authentizität in der Briefedition: Integrale Darstellung nichtsprachlicher Informationen des Originals." *editio* 16: 36–56.

NOTES

- 1 The term "letter" in the following means "piece of correspondence" in general.
- 2 <http://wyatt.elasticbeanstalk.com/mep/home>
- 3 See David Chesnutt, C. M. Sperberg-McQueen, and Susan Hockey, "TEI/MEP Encoding Scheme," version 4.0, January 16, 2002, <http://wyatt.elasticbeanstalk.com/mep/misc/mepw04.html>, and David R. Chesnutt, Susan M. Hockey, and C. M. Sperberg-McQueen, "Markup Guidelines for Documentary Editions," July 4, 1999, <http://wyatt.elasticbeanstalk.com/mep/misc/mepguide.xmlver.html>.

Although no explicit namespace is used in the MEP project, we nevertheless prefix element names under discussion (in running text) for better comprehensibility. Elements without namespace prefix are from the official TEI standard.

- 4 Centrum voor Teksteditie en Bronnenstudie, Koninklijke Academie voor Nederlandse Taal- en Letterkunde (Centre for Scholarly Editing and Document Studies, Royal Academy of Dutch Language and Literature), <http://ctb.kantl.be/project/dalf/index.htm>.

- 5 Edited by Leo Jansen, Hans Luijten, and Nienke Bakker, 2009, <http://vangoghletters.org/>. This project adapted some elements needed for the encoding of letters from the DALF project, such as the element <vg:letDesc> for bundling together the core metadata of the letters: see http://vangoghletters.org/vg/about_6.html#intro.VI.6.1.
- 6 Feature request by Nick Finke, May 20, 2004, closed March 2, 2007, Text Encoding Initiative SourceForge site, <http://sourceforge.net/p/tei/feature-requests/3/>.
- 7 Joseph Jung, ed., *Digitale Briefedition Alfred Escher* (rev. ed. Zürich: Alfred Escher-Stiftung, July 2015), <http://www.briefedition.alfred-escher.ch>.
- 8 Maurizio Ghelardi et al., *Burckhardtsource Project* (The European Correspondence to Jacob Burckhardt) (2015–), accessed March 18, 2016, <http://burckhardtsource.org>.
- 9 Peter Stadler and Joachim Veit, eds., *Carl-Maria-von-Weber-Gesamtausgabe* (2003–), accessed March 18, 2016, <http://www.weber-gesamtausgabe.de/en/>.
- 10 Anne Baillet, ed., *Letters and Texts. Intellectual Berlin around 1800*. (2012–), accessed July 25, 2016, <http://www.berliner-intellektuelle.eu/?en>.
- 11 Susan Schreibman et al. (2013–), accessed March 18, 2016, <http://letters1916.ie/>.
- 12 Ed Folsom and Kenneth M. Price, eds., Center for Digital Research in the Humanities, University of Nebraska–Lincoln (1995–), accessed March 18, 2016, <http://www.whitmanarchive.org/>.
- 13 Buffalo Bill Center of the West and University of Nebraska-Lincoln, accessed March 18, 2016, <http://codyarchive.org/>.
- 14 *Digitale Edition der Korrespondenz August Wilhelm Schlegels*, accessed March 21, 2016, <http://august-wilhelm-schlegel.de/>.
- 15 Andrew Jewell, Janis P. Stout, and Melissa J. Homestead, eds., Center for Digital Research in the Humanities at the University of Nebraska–Lincoln, last modified February 2016, <http://cather.unl.edu/completelethers.html>.
- 16 *Epistolary Networks: Visualising Multi-dimensional Information Structures in Correspondence Corpora* (Vernetzte Korrespondenzen: Erforschung und Visualisierung sozialer, räumlicher, zeitlicher und thematischer Netze in Briefkorpora), Trier Center for Digital Humanities, accessed March 21, 2016, <http://exilnetz33.de/en/>.
- 17 Berlin-Brandenburg Academy of Sciences and Humanities, accessed March 24, 2016, <http://correspsearch.bbaw.de/>.

- 18** Cultures of Knowledge Project, Bodleian Library, University of Oxford, accessed March 21, 2016, <http://emlo.bodleian.ox.ac.uk/>.
- 19** Stanford University, accessed March 21, 2016, <http://republicofletters.stanford.edu/>. This project and EMLO (see [note 18](#)) both work with linked open data in which TEI is included but not the only possibility.
- 20** Visual Correspondence: Analysing Letters through Data Visualisation, Niall O’Leary Services, accessed March 21, 2016, <http://letters.nialloleary.ie/>.
- 21** For more digital correspondence editions, see the catalogue compiled by Patrick Sahle (*A Catalog of Digital Scholarly Editions. Material: Letters*, Institute for Documentology and Scholarly Editing, accessed March 21, 2016, http://www.digitale-edition.de/vlet_letters.html), or the Correspondence SIG’s wiki page (last modified December 6, 2015, <http://wiki.tei-c.org/index.php/SIG:Correspondence>). There are also correspondence editions that decided against using the TEI, like the project *The Diplomatic Correspondence of Thomas Bodley, 1585-1597* (version 5, edited by Robyn Adams, Centre for Editing Lives and Letters, 2011, <http://www.livesandletters.ac.uk/bodley/encoding.html>). However, the possibility of transforming all XML files into a format compatible with TEI after the completion of the project was kept open.
- 22** <http://www.kantl.be/ctb/project/dalf/>.
- 23** “ODD files for documenting the Digital Edition of the Carl-Maria-von-Weber-Gesamtausgabe,” accessed March 22, 2016, <https://github.com/Edirom/WeGA-ODD>.
- 24** Definition of the element in [Van den Branden 2013](#), file DALF .htm.
- 25** Definition of the element in [Van den Branden 2013](#), file DALF .htm.
- 26** German original version: “Wichtiger als die Frage, was sich ereignet, ist die Tatsache, daß sich etwas ereignet, also der bloße Vollzug.” Translation by the authors.
- 27** German original version: “Die ... Begleitumstände – übersandte Schriften, Geschenke, Kunstwerke, weitere, mündliche Nachrichten – zeigen, daß der Brief in ein ganzes Ensemble aus Kommunikationsmedien eingebunden war, welche die Korrespondenz zu einem multimedialen Prozeß machten.” Translation by the authors.
- 28** <http://www.tei-c.org/Vault/P5/2.8.0/doc/tei-p5-doc/en/html/ref-author.html>.
- 29** <http://www.tei-c.org/Vault/P5/2.8.0/doc/tei-p5-doc/en/html/ref-profileDesc.html>.
- 30** <http://www.tei-c.org/Vault/P5/2.8.0/doc/tei-p5-doc/en/html/ref-correspDesc.html>.

- 31 <http://www.tei-c.org/Vault/P5/2.8.0/doc/tei-p5-doc/en/html/ref-correspAction.html>.
- 32 <http://www.tei-c.org/Vault/P5/current/doc/tei-p5-doc/en/html/ref-correspContext.html>.
- 33 <http://www.tei-c.org/Vault/P5/2.8.0/doc/tei-p5-doc/en/html/CC.html#CCAS2>.
- 34 Last modified March 19, 2015, https://github.com/TEI-Correspondence-SIG/correspDesc/blob/master/examples/example07_anthology-of-letters.xml.
- 35 “A Proposal for a Correspondence Module and a correspDesc Element,” last modified March 20, 2015, <https://github.com/TEI-Correspondence-SIG/correspDesc/>.
- 36 Deutsche Nationalbibliothek, last modified February 29, 2016, <http://www.dnb.de/EN/Standardisierung/GND/gnd.html>.
- 37 OCLC, accessed March 24, 2016, <http://viaf.org/>.
- 38 Misha Wolf and Charles Wicksteed, “Date and Time Formats,” W3 Consortium, September 15, 1997, <http://www.w3.org/TR/NOTE-datetime>.
- 39 <http://correspsearch.net/>.
- 40 Syd Bauman, comment on TEI SourceForge feature request no. 3, “Letters/memos module,” March 2, 2007, <https://sourceforge.net/p/tei/feature-requests/3/>.
- 41 SIG:Correspondence/ODD work, TEI Wiki, http://wiki.tei-c.org/index.php/SIG:Correspondence/ODD_work. The “first draft” was by Peter Stadler (Stadler 2011).
- 42 Cf. Note 34.
- 43 The Manuscripts SIG workgroup on Genetic Editions could serve as a role model here. They held two workshops refining their draft proposal, described by Elena Pierazzo on her blog (“At the TEI Technical Council: Genetic Criticism Encoding,” *Elena Pierazzo’s Blog*, November 8, 2011, <http://epierazzo.blogspot.de/2011/11/at-tei-technical-council-genetic.html>).
- 44 Peter Stadler, feature request #510: “Add a Correspondence Module and Elements for Capturing Correspondence Specific Meta Data,” last modified May 20, 2015, <https://sourceforge.net/p/tei/feature-requests/510/>.

ABSTRACT

The encoding of letters has a long tradition in the TEI but there have been no official recommendations in the TEI Guidelines on how to deal with correspondence. Two TEI customizations present exemplary models: *DALF: Digital Archive of Letters in Flanders* and *Carl Maria von Weber—Collected Works (WeGA)*. These were the basis of the work of the TEI Correspondence Special Interest Group that formed a task force—consisting of the three authors of this article—for developing encoding guidelines. This article discusses correspondence theory in brief, letters as an act of communication, and how these aspects of correspondence can be expressed by TEI entities. The development of the communication-oriented concept of correspondence and its direct implementation in correspondence-specific metadata structures will be discussed. Central is the new wrapper element <correspDesc> (correspondence description), which stores key metadata about the encoded piece of correspondence. After addressing this first question of how one should encode correspondence with the TEI, we will discuss the question of linking and interchange between projects and editions dealing with correspondence material. To facilitate this, the Correspondence Metadata Interchange (CMI) format has been developed by the TEI Correspondence SIG’s task force as a subset of <correspDesc>. Finally, we will describe some organizational difficulties in implementing the new elements and encoding model into the TEI Guidelines in 2015.

INDEX

Keywords: correspondence, communication, digital edition, interchange, metadata, authority files

AUTHORS

PETER STADLER

Peter Stadler is a researcher at *Carl Maria von Weber—Collected Works* at the University of Paderborn, Germany, where he is in charge of the digital edition of Weber’s letters, diaries, and writings. He was the co-initiator (in 2008) and is currently co-convenor of the TEI Correspondence SIG and an elected member of the TEI Technical Council for the term 2014–15.

MARCEL ILLETSCHKO

Marcel Illetschko holds a PhD in Modern German Literature from the University of Vienna (2011) and is one of the conveners of the TEI Correspondence SIG. He is a researcher at the Literature Archive of the Austrian National Library (working on *The letters of August Sauer and Bernhard Seuffert*, Austrian Science Fund/FWF: I730-G20) and also works at the Federal Institute for Education Research, Innovation & Development of the Austrian School System (Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens, BIFIE).

SABINE SEIFERT

Sabine Seifert is a PhD candidate in Modern German Literature at Humboldt University Berlin and is one of the conveners of the TEI Correspondence SIG. She is a researcher at the Theodor Fontane Archive at Potsdam University. From 2010 to 2015, she worked at the Berlin Intellectuals 1800–1830 project at Humboldt University Berlin, where she contributed to the development of the digital edition *Letters and Texts: Intellectual Berlin around 1800*.