TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

Anne Baillot and Julie Giovacchini

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward
a Standardized Workflow

1

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

Anne Baillot and Julie Giovacchini

ABSTRACT

The TEI Guidelines are developed and curated by a community whose main purpose is to standardize the encoding of primary sources relevant for humanities research and teaching. But other communities are also working with TEI-based publication formats. The first goal of this paper is to raise awareness of the importance of TEI-based scholarly publishing as we know it today. The

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

2

second goal is to contribute to a reflection on the development of a TEI customization that would cover the whole authoring-reviewing-publishing workflow and guarantee archiving options that are as solid for journal publications as what we now have for primary sources published in TEI.

INDEX

Keywords: workflow, digital publications, schemas, scholarly publishing

1   The TEI Guidelines are developed and curated by a community whose main purpose is to standardize the encoding of primary sources relevant for humanities research and teaching. This common focus on a research and dissemination goal, which is rooted in the history of the TEI, gives the community a great strength, but it leads to a lack of recognition for other areas of digital textuality that also use the TEI. We mostly encode primary sources relevant to our research and concentrate development efforts in their direction. But other communities are now also working with TEI-based publication formats.

2   Looking at the number of encoded documents, the production of the TEI community as described above (e.g., actively developing the guidelines and doing so with a focus on primary sources) is by far not the most productive among those who currently use TEI. The best-known example for this uneven situation is that of the European Patent Office: all its patents are encoded in TEI, amounting to a total of two hundred million documents and two billion TEI annotations (Laurent Romary, personal communication). This kind of magnitude is far from that achieved by all TEI-based humanities primary sources put together.

3   A lack of balance also exists within the humanities ecosystem, where the TEI is used to encode not only primary sources, but also secondary sources, especially journal publications. In that context, the quantity of TEI-based documents is quite impressive without there being much theoretical or research-oriented development made in the field of journal publications, even though it is a growing one. TEI publication formats for secondary sources are evolving and being developed, but rather like a by-product. They are seldom addressed on TEI-L, at the Annual Conference, or in the *Journal of the Text Encoding Initiative.* The latest example of such a publication, for instance, addresses OJS-based workflow issues, but not the encoding per se (Homenda and Pekala 2016).

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

3

4   This paper draws on the questions that were at the core of Holmes and Romary (2011) and that initiated the creation of the jTEI format, now used among others by the *Journal of the Text Encoding Initiative*.[1] It is not our purpose to discuss this schema or possible amendments to it, nor to discuss the respective technical benefits of TEI vs. JATS (Journal Article Tag Suite, http://jats.niso.org/1.1/). We have two distinct goals in building upon Holmes and Romary (2011). The first goal is to raise awareness of the importance of TEI-based scholarly publishing as we know it today. The second is to contribute to a reflection on the development of a TEI customization that would cover the whole authoring-reviewing-publishing workflow and guarantee archiving options that are as solid for journal publications as what we now have for primary sources published in TEI. The encoding ideas we propose are to be considered as explorative.

5   In a first step, we will present a panorama of the use of TEI in social sciences and humanities (SSH) scholarly publishing and consider the advantages and challenges of using TEI-based formats in scholarly publishing in the humanities. Which organs are currently doing so and how much text does this represent? Why are these journals using a TEI format? We will then relate our experience as editors, with the main aim of initiating a discussion within the community on the role of TEI-based formats for scholarly publishing at large, focusing on the copyediting and reviewing process. We will argue that such formats have the potential to be a powerful leverage to increase the TEI's impact on the scholarly community and to empower humanities scholars for better dissemination of their own research. We will propose some preliminary ideas for encoding a complete publishing workflow of secondary sources that include the reviewing and copyediting process.

1. Scholarly Publishing in TEI: An Overview

6   Scholarly publishing practices are in many ways connected to the existing research infrastructures that are available to researchers and librarians. The development of TEI-based infrastructures and services on different continents and in different countries is the result of personal engagement by members of the TEI community and of the historical development of infrastructures in these different environments. It also has an economic dimension (who is paying how much for which service) that impacts greatly the digital publishing landscape as we know it today.

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward
a Standardized Workflow

4

7   The following is an overview of TEI-based journal publications that was gathered mainly by initiating an informal survey on the TEI-L discussion list.[2] Since there is no other overview of this type known to us in this area, we assume it is the best summary of the current situation.[3]

1.1 In Europe

8   The largest actor in Europe is France, where a long history of political centralization led to the development of national infrastructures and services that has allowed the deployment of TEI-based publishing formats and TEI-compliant platforms on a national scale. France has three main providers for secondary sources that rely on the TEI: OpenEdition, HAL, and Istex.

9   OpenEdition[4] is a platform hosting four different services: an academic calendar announcing events such as conferences and call for papers (Calenda), a scholarly blog platform called Hypotheses, a scholarly book service (OpenEdition Books), and a scholarly journals platform (OpenEdition Journals, called revues.org in the past). The *Journal of the Text Encoding Initiative* is hosted by OpenEdition. Almost all of the journals published on the journals.openedition platform are TEI-based.[5] The articles published as of February 15, 2021 amount to a total of 310,515 documents.[6] The book platform also provides many TEI-based documents (204,869 as of February 15, 2021),[7] which adds up to a total of over half a million documents, some of them being books, that is, rather large documents.

10   Although the TEI files for these articles and books are theoretically available for reuse and can be used for research purposes, there is currently no direct access like a download link or button that would make harvesting easy. When asked about the feasibility of such an endeavor, OpenEdition mentions[8] local legal issues but a clear willingness to support research projects that would need to download these files, should such a research project issue a request (which has never yet ever happened).

11   The second French infrastructure to be TEI-based is not strictly a platform for scholarly publishing but a publication repository that can be used for either preprint or postprint open access (OA) publication. Unlike OpenEdition, which was initiated by a research project and gained momentum because it met the needs of the scholarly community and the national political agenda, the open archive HAL[9] was from its onset a national infrastructure, conceived as a service to the French scholarly community at large, with a specific effort directed towards humanities disciplines that

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

5

had resisted noncommercial archiving strategies even long after other communities relied on them. Scholars can either register metadata concerning their publication or additionally archive one or several preprint versions. It is the metadata that are stored in a TEI format, potentially allowing a wide overview of French scholarly publications.

12 The significance of the HAL data has grown since July 2018 when the French Ministry for Higher Education and Research issued an Open Science Plan that led the main funding agency ANR to require HAL publications of research output they fund from then on.[10] It was followed in this constraining requirement by the national research evaluation organ HCERES, which issued a statement that they would only take into consideration full-text HAL publications when evaluating universities.[11] As for the CNRS (a national organization employing only researchers and research assistants) and Inria (national organization for research in informatics), they also require researchers to use HAL for evaluation purposes on both the lab and the individual level. At the CNRS, HAL import functionalities are currently being transferred into the internal databases that harvest annual reports.[12] Moving to the TEI-based HAL is a political evolution that took some time to be implemented throughout the research ecosystem, but as of 2020 there is virtually no single publication produced in France that will not have at the very least a set of TEI metadata associated with it. The centralization of information which these evaluation and archiving constraints have imposed on a wide range of scholarly communities has not always been well received. In the humanities and social sciences, it was occasionally considered[13] to have the potential for enabling political control, and hence to be detrimental to the freedom of research. This was especially a concern for those who were not familiar with the underlying technologies or the values of the community that develops them.

13 The third French publication organ hosting TEI-based information is ISTEX.[14] ISTEX is a platform that integrates content from private scholarly publishers (OpenEdition being one of them) with whom national licenses are being negotiated. University libraries usually contribute a yearly amount to benefit from the deal, which is managed at a national level by a negotiation consortium (Consortium Couperin, https://www.couperin.org/). ISTEX contains metadata from publications, but also full-text PDF files, that are being transformed into TEI and added to the server. ISTEX' API is used by university libraries to retrieve information as well as for bibliometric purposes. ISTEX currently hosts twenty-three million documents.[15]

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

6

14   All of this together amounts to a considerable mass of documents, but not all of them are available in TEI and not all of them are available in full OA from the onset. There still are accessibility restrictions, whereas in the case of the examples we gathered from other countries, all journals are completely TEI-based (including both metadata and text) and OA.

15   Germany has a different strategic approach, but it still presents a series of initiatives in the field of TEI-based scholarly publishing. Until recently, Germany relied solely on the good will (political and economic) of its regions, and more specifically of their universities, to build and maintain the infrastructures that are necessary for hosting reliable publication platforms. A funding scheme for a national infrastructure has recently been negotiated for hosting scholarly data at large in a stable environment.[16] Building such an infrastructure will certainly change the overall approach of scholarly publishing in the middle and long run. At this stage though, it has not been fully implemented and we still are in a rather fragmented situation, despite the coordination efforts provided by the DARIAH[17] and CLARIN[18] research infrastructures (both are European Research Infrastructure Consortiums, or ERICs).[19]

16   Several academic actors in the social sciences and humanities are positioned in the field of TEI-based scholarly publishing. The *Zeitschrift für digitale Geisteswissenschaften*[20] was created together with the German-speaking branch of the Alliance of Digital Humanities Organizations (ADHO), DHd (Digital Humanities im deutschsprachigen Raum). It focuses on digital humanities at large, and the articles are mostly in German. Eighty-one articles are currently published, all in TEI P5.

17   R-I-D-E[21] does not follow a similar institutional logic in the sense that it is not carried by one scholarly society, but by the virtual Institute for Documentology and Edition sciences, the I-D-E.[22] After the institute concentrated for several years on teaching and research, its publication organ R-I-D-E has gained momentum over the past several years. All of their publications are TEI-based and the data can be retrieved easily.[23]

18   Finally, the *Jahrbuch für historische Bildungsforschung*[24]—a publication backed by a solid institution, in this case, the Leibniz institute for the History of Education —has also been preparing to convert to a full TEI-based, OA publication in an eXist database. While it will use the jTEI Article schema for the scholarly journal, the goal is to generate continuity between other scholarly resources that will also be published in TEI, and the journal itself.

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward
a Standardized Workflow

7

19   Elsewhere in Europe, Slovenia has one journal that is TEI-based: *Prispevki za novejšo zgodovino /
Contributions to Contemporary History,[25] following the jTEI Article schema. The journal has published
181 articles as of 2019.

1.2 In North America

20   Among the scholarly journals based on a TEI format, *Digital Humanities Quarterly* is the most
prominent. It has strong ties to the TEI community (its editorial board including several active
members of the TEI community) and uses a customization of TEI P5. *Scandinavian-Canadian Studies*,
which had published 185 articles and reviews as of March 21, 2021, uses P4 converted to P5. Indiana
University has two TEI-based journals: *Indiana Magazine of History* (IMH) and *The Medieval Review*
(TMR) (Homenda and Pekala 2016). The journal of Beth Mardutho: The Syriac Institute, *Hugoye:
Journal of Syriac Studies*, has recently been migrated to TEI and is currently converting all back issues
to TEI.[26] *Romantic Circles* (University of Colorado Boulder) has been working with TEI since 2011.[27]

2. Advantages and Challenges of using a TEI-based Format
for Scholarly Publishing

21   Publishing scholarly journals in TEI offers several advantages on different levels.

22   On an economic level, using the jTEI Article schema on top of an OJS workflow allows low-budget
production of scholarly publications like journals. OJS is an open-source platform that can be
customized and maintained with the support of an active user community. The jTEI Article schema
is available for any journal to reuse, and there exist transformation scenarios for the OpenEdition
platforms. In this process, the development needed is limited. In terms of output, the result is
clean, and in case of problems or bugs, the editors can mostly rely on the community to tweak
minor issues. In other words, you do not need to invest a lot to publish your journal using such an
infrastructure and in return, there is little economic gain to be expected from it.

23   The low cost and the easy transfer to open science allowed by such an infrastructure can appeal to
decision makers like university presidents, especially for scholarly domains that are comparatively
not too impact factor–dependent. There is, on a more political level, a growing interest in stable,
sustainable publishing solutions that are now increasingly being recognized as economically
valuable.[28]

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

8

24 These arguments can explain why convincing decision makers, and more generally people from outside the TEI community, to "invest" in a TEI-based workflow for scholarly publishing has become a worthwhile effort, especially in times when the values of open science—a philosophy with which the TEI technology is compatible—are being promoted.

25 We base this assessment on the existence of still-isolated but real attempts, which show the ability of TEI schemas to propose a satisfactory processing that complies with the main principles of accessibility and interoperability[29] of academic data and metadata.[30]

26 Two major needs can indeed be identified in the context of SSH journals. The first is the need for an interoperable and stable workflow that would be integrated to open source publication infrastructures. The second need is for an evaluation and publication process that would include transparency as a core virtue. While our proposition, described in more detail in the third part of this paper, is aiming at massively improving interoperability and stability by relying on certain characteristics of TEI particularly adapted to these problems (use of a controlled vocabulary, possibilities of semantic and descriptive markup), it is also our goal to encourage discussion of transparency and interoperability as key quality criteria in SSH journals. In that sense, the use of TEI technology in an editorial process is likely to affect not only workflows but also evaluation criteria in general. It is indeed intended to increase the consideration for editorial tasks in the context of Open Science in particular, but also of SSH research in general.

27 For the TEI community, pushing forward secondary sources in TEI (as opposed to primary sources) presents several advantages. The first aspect worth mentioning is an assessment of current realities such as was provided in the first part of this paper: we have reached more than a critical mass already, one that calls for better coordination and sustainability of how the community integrates its outputs.

28 Why does TEI encoding work so well for scholarly publishing of secondary sources? There are several possible reasons. First, the TEI is flexible in its vocabulary, which means that it allows us to manage and bring together heterogeneous sources of information. Second, the TEI is not closed. On the contrary, it is conceived so as to allow resources to communicate: it enables us to avoid silos of internal formats developed for the use of one specific publication platform that will need further specific development to remain sustainable, costing a lot to be maintained and stay in use. What makes the TEI valuable for secondary scholarly publications is what makes it valuable for

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

9

any publication: its stability, its interoperability, its openness, its reusability. TEI-based scholarly publications are made available in a nonproprietary format, which can also easily be transferred to OA publication models.

29   There is one final argument, proposed by Laurent Romary in the course of the TEI-L discussion on the topic, that is less obvious than the previous ones. To him, the main advantage of the TEI being the same basis format for primary and secondary sources is that it allows scholars to use the same format for primary and secondary scholarly publications: for example, for digital editions and articles on the digital editions. This continuity between the two major publication dimensions of SSH research activity allows both fluidity and solidity. Fluidity means here that integrating elements from primary resources to secondary resources (and reciprocally) is made particularly easy. Solidity means that the same people who have the expertise in one field can contribute to the other: the brainpower available is considerable. But it also means that issues of nesting TEI structures are of central importance in this context.

30   In terms of research content, the continuum between primary and secondary scholarly publications could have another consequence: the TEI community could (or should?) become more attractive for scholars interested in less philological questions than the core community is: for instance, sociological aspects of knowledge transfer, community building, or evolution of research strategies. Journal material could easily be turned into a primary source. And there are enough documents available at this point to make this corpus interesting to sociologists, sociolinguists, and historians of science. What is still missing most of the time is a more obvious download button —that is, a structural incentive to use TEI corpora.

31   The advantages of expanding TEI-based models to secondary publications would not only benefit the TEI community by widening it. It would also make it possible to bypass the dead end of reputation mechanisms in the SSH at large. We are currently in a situation where scholars, assistants, and research engineers are sacrificing a great deal of time and work for the profit of publishing houses that are negotiating reputation for money. Coming up with a TEI-based format has the potential to break these reputation rules because it is nonprofit, requires little technical maintenance, and allows scholars to dedicate more time to reading papers than evaluating their impact factor or abiding by publishers' editorial guidelines (see Kosmopoulos and Pumain 2008).

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward
a Standardized Workflow

10

32   What are TEI-based scholarly publications aiming at in general? Mainly, stable and wide dissemination. What makes reputation in the realm of such values is neither primarily quantity nor established publishing houses but mostly a dissemination strategy based on core virtues like FAIR (Findability, Accessibility, Interoperability, and Reuse),[31] and a common set of values and improvements carried out by the community. Researchers need relevant papers to be accessible and easy to find, to evaluate and to reuse; TEI-based formats can help meet these needs. They are already doing so for publishing in the situations described in part 1, and sometimes also for authoring, but never so far for reviewing and copyediting.

3. TEI-based Workflow Improvements

33   To be able to provide a complete publication workflow in TEI, we would not only need to improve and develop customizations for publishing and for authoring as they already exist.[32] We need to conceive a complete TEI-based workflow, to take into account the reviewing and copyediting phase as well—phases that are currently dealt with in other formats.

34   Complex workflows are at the core of editorial work for scholarly journals. The coordination chain usually includes authors, editors, reviewers, and copyeditors. Depending on the journal and the reviewing format (single-blind, double-blind, open), this can easily add up to ten people working on one text. Our own editorial experience, respectively at the *Journal of the Text Encoding Initiative* and *Philosophie Antique*, can give significant insight into such workflows. *JTEI*, for instance, foresees three different reviewers for each paper and three rounds of copyediting (some of which are done by the editors, but not all). *Philosophie Antique*[33] has a printed edition in addition to the digital one; this implies, in addition to double-blind review for each article, a double copyedit for each of the two formats. With different issues running in parallel as is now the case in most online publications, this means having an editorial interface that makes it possible to deal with different workflow timelines at the same time and to assign different editorial roles to one single person.

35   *JTEI* is working with OJS, as are many TEI-based journals. The recent update to version 3.0 has made some improvements to OJS's functionalities and interface, but OJS still suffers from being developed for too many different uses, making it occasionally tricky to tailor for specific needs. For years, the overall management of *JTEI* was actually dealt with not in OJS itself, but in tables archived in separate Google documents (one Google document for each issue of the journal), because OJS

proved unable to offer such an overview in the way that was needed. This type of management has the inconvenience that the text and the information on the status of the text are separated, making an overview difficult to gain for the different actors involved in the process of text production. *Philosophie Antique* is even worse off. The workflow does not use any content management interface. The editorial team works by document exchange, archiving successive versions, with two parallel workflows for the preparation of the PDF to be printed and the uploading of texts in HTML, using the Lodel tool on the OpenEdition interface.

36 These two examples are symptomatic of the situation in many journals. Compared to this reality, it is clear that, to be able to deal with the complexity of workflows as we know them today, the texts should ideally contain metadata allowing one to see at first glance, as well as to process automatically, their editorial status. This is one of the many improvements that the TEI can provide.

37 To conceive a TEI encoding that would fulfill the reviewing and copyediting function and allow for an overview of the editorial status of the text, let us first consider the different roles (author, reviewer, copyeditor) and the different types of interventions in the text. In an ideal world, each role would correspond to a type of intervention, but in actual editorial practice, it might well be that the copyeditor who checks for content coherence also finds typos to correct. One can divide roughly into two categories the types of interventions that will be done on the text: one intervention type encompasses content (editorial requirements and their application), while the other one deals with the form (ortho-typographical aspects). This corresponds to two workflows that run in parallel: one checking for the scholarship, adequacy, and coherence of the article, and one checking for its conformance to the typographical guidelines of the journal. The TEI encoding for reviewing and copyediting should reflect these two aspects as well as the two work phases that are the reviewing and the copyediting processes.

38 The first step in a TEI-based workflow will be to define on the one hand roles (e.g., reviewer A, reviewer B, and reviewer C) and on the other hand types of modifications that can be undertaken: an editorial schema (EdSchema), a tagset for modifications of formal aspects, and a redactorial schema (RedSchema) that allows the tagging of content-related modifications. To illustrate the

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

12

prototype we have in mind, we used existing examples of articles already published or in the process of being published and converted the copyediting/reviewing process from the Word document to a TEI-based version.[34]

Figure 1. Workflow pattern.
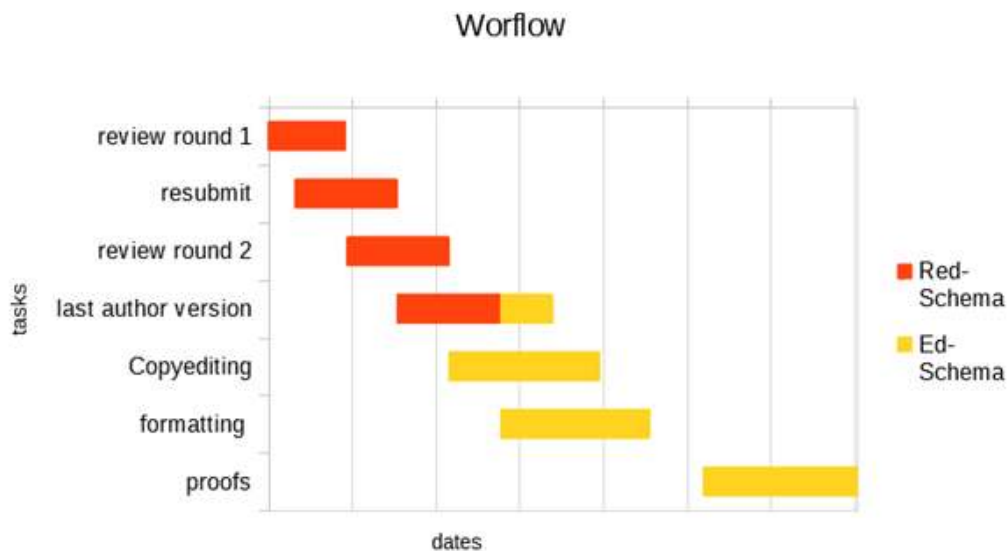
39   The graph in figure 1 is a simplified representation of the annual workflow of the journal *Philosophie Antique*. Each article submitted and published by the journal goes through the same process.

40   The first part of the process is dedicated to the scholarly evaluation of the paper's content; it concerns the reviewers (whose names remain unknown to the authors), the authors, and the journal's editors. All this part would be encoded using the `RedSchema` subscheme.

41   The work of formal preparation for publication takes place in a second stage that partially overlaps the first one. It partly concerns the authors (at two different moments: delivery of the last version of the text, and proofreading) and the editorial team. It would be encoded using the `EdSchema` subscheme.

42   We find a strong advantage to this visualization of the workflow by task cycle and not by agent because it allows us to de-individualize the different interventions on the text, which must essentially be approached not according to their source (who intervenes) but according to their nature (what type of intervention). The actors within each stage are then only differentiated by a

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

13

@resp attribute, which can be anonymized according to the editorial needs (especially in the first stage for a double-blind review), and their interventions fit into one or the other of the subschemes depending on whether they concern content, or form and presentation.

43   EdSchema allows the tagging of elements from the review process as well as from the copyediting process. The tags are attributed to the different <resp>s defined in the <respStmt> part of the header and associated with an @xml:id (see figures 2 and 3 and figures 10, 11, and 12). EdSchema contains primarily the <lem>, <add>, <del>, and <choice> elements.

Figure 2. Review process with track changes.

> In addition to the reference format, each collection is made available in a simple plain-text
> format automatically derived from the TEI~~XML TEI~~ version, containing only the text included
> in the body of the narrative texts and plays (in particular, excluding prefaces and other paratext
> as well as notes) and with external metadata provided in tabular format. This format is especially
> suitable ~~notably~~ for direct use with the stylo ~~tool~~ package ~~stylo~~ for R (Eder, Kestemont, and
> Rybicki ~~et al.~~ 2016) and other tools operating on the surface level of texts.

Figure 3. Tags from EdSchema used for the review process.

```
In addition to the reference format, each collection is made available in a simple plain <del resp="#Ed1">-</del>text
format automatically derived from the <del resp="#Ed1">XML</del>TEI version,
containing only the text included in the body of the narrative texts and plays
(in particular, excluding prefaces and other paratext as well as notes)
and with external metadata provided in tabular format. This format is
   <choice> <sic resp="#aut">suitable notably</sic><corr resp="#Ed1">especially suitable</corr> </choice>for direct use with the stylo
   <choice> <sic resp="#aut"> tool</sic> <corr resp="#Ed1">package</corr></choice> for R
   (Eder <choice><sic resp="#aut">et al.</sic><corr resp="#Ed1">, Kestemont, and Rybicki</corr></choice> 2016)
and other tools operating on the surface level of texts.
```

44   RedSchema consists of a tagset that is in part similar to that of EdSchema, with each tag attributed to its corresponding role (author, reviewer, editor ; see figures 4 and 5).

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

14

Figure 4. Scholarly evaluation and comments in a word processor.
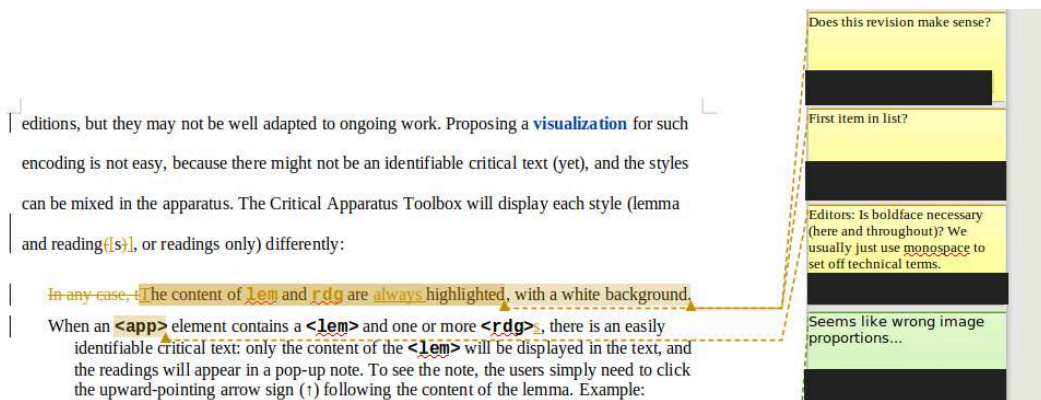
Figure 5. Scholarly evaluation and comments with `RedSchema`.

```
<app from="#Rev1beg" to="#Rev1end" resp="#Ed1"> <note> Does this revision make sense? </note></app>
    <app from="#Rev2beg" to="#Rev2end" resp="#Ed1"> <note>First item in list?</note></app>
    <app from="#Rev3beg" to="#Rev3end" resp="#Ed1"> <note>Editors: Is boldface necessary (here and throughout)?
        We usually just use monospace to set off technical terms.</note></app>
```

45 Both types of interventions are likely to involve short alterations (changes in punctuation marks, for instance), bibliographical elements (see figures 8 and 9),[35] and longer text passages that need re-writing, this last category being more likely to be relevant to `RedSchema`.

46 To address remarks that require the rewriting of a longer text passage, `RedSchema` needs to include an anchor-based tagset that allows pointing to a comment, which in turns should allow the author(s) and/or editor(s) to answer this comment (see figures 6 and 7, figures 8 and 9, figures 10, 11, and 12 and figures 13 and 14).

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

15

Figure 6. Copyediting discussion in a word processor.

for the current status quo in the TEI Guidelines or directly influenced our own understanding of the gGuidelines and how they should evolve.

A major milestone in the digital dictionary era is the work by Amsler and Tompa (1988), which, together with the unifying contributions of Ide and Véronis (1994), led to the earlier TEI "Print Dictionaries" chapter. Focusing here on their contribution to etymology, example 1 shows how they have introduced a highly structured model based on etymons and links implemented as an SGML DTD. The underlying model is based upon a graph of etymons (`<etymon>`) connected with relations (`<rel>`), forming a more global etymological tree (reflected by the

Is this the right term?

"earlier" seems fine to me.

"early"?

This would be a good place for a specific citation to that earlier chapter.
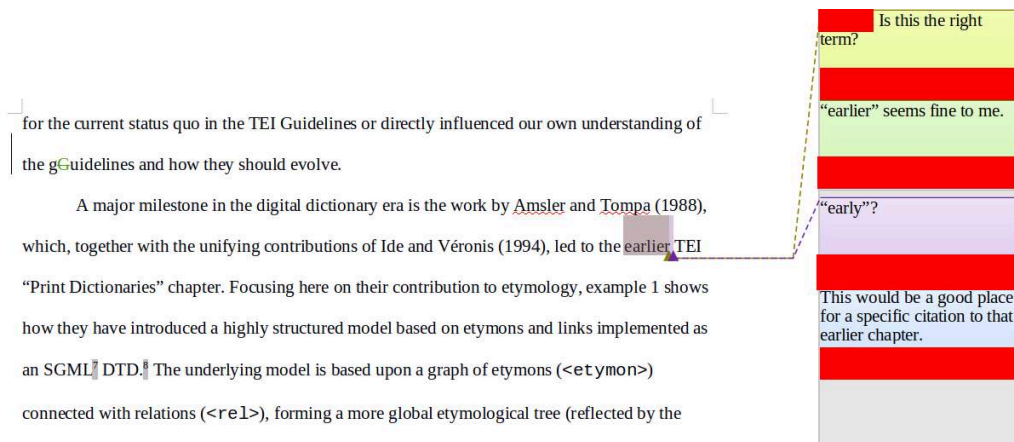
Figure 7. `RedSchema` with discussion.

```
<div type="texte" resp="#aut">
<p>A major milestone in the digital dictionary era is the work by Amsler and Tompa (1988),
    which, together with the unifying contributions of Ide and Véronis (1994),
    led to the <anchor xml:id="Rev1beg"/>earlier<anchor xml:id="Rev1end"/> TEI "Print Dictionaries" chapter.
    Focusing here on their contribution to etymology, example 1 shows
    how they have introduced a highly structured model based on etymons
    and links implemented as an SGML  DTD.  The underlying model is based upon a graph of etymons (@etymon)
    connected with relations (@rel), forming a more global etymological tree
    (reflected by the @es, etymological segment, element).</p>
</div>

<app xml:id="Rev1" from="#Rev1beg" to="#Rev1end"><note resp="#Ed1">    : Is this the right term?</note>
    <note resp="#Ed2">"earlier" seems fine to me.</note>
<note resp="#Ed3">This would be a good place for a specific citation to that earlier chapter.</note></app>

<div type="discussion">
    <p resp="#aut"> <ptr target="#Rev1"/>"early"?</p>
</div>
```

Figure 8. `RedSchema` with bibliographical improvement from reviewer.

```
<p>Mais Lucrèce, dans un geste poétique que cette intertextualité révèle donc comme hautement <anchor xml:id="Rev1beg"/>
subversif<anchor xml:id="Rev1end"/>, choisit de transformer cette situation dangereuse et immorale en une relation
paisible, joyeuse et bienfaisante — puisque l'union de Mars et de Vénus représente l'apaisement de la brutalité et de la
sauvagerie par la recherche du plaisir selon la nature.</p>

<app xml:id="Rev1" from="#Rev1beg" to="#Rev1end"> <note resp="#Ed1">Pour nuancer un peu le caractère subversif de ce
geste, voir <bibl >Françoise Gury, 'Mars et Vénus : l'amour dans la dignité : le dossier de la peinture campanienne', in
De la théâtralité du corps aux corps des dieux dans l'Antiquité, ed. by Valérie Huet and Florence Gherchanoc (Brest:
Centre de Recherche Bretonne et Celtique, 2014), pp. 143-161</bibl></note></app>
```

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

16

Figure 9. Comment in word processor with bibliographical improvement from reviewer.

mais il l'associe discrètement à son époux légitime¹. Mais Lucrèce, dans un geste poétique que cette intertextualité révèle donc comme hautement subversif, choisit de transformer cette situation dangereuse et immorale en une relation paisible, joyeuse et bienfaisante – puisque l'union de Mars et de Vénus représente l'apaisement de la brutalité et de la sauvagerie par la recherche du plaisir selon la nature.

Ce choix poétique place dès l'incipit la sexualité sous le signe de l'ambiguïté ; elle est une nécessité non seulement physique mais éthique, elle est également un piège dont il faut se méfier – piège au sens propre du terme, la scène dépeinte par Lucrèce précédant immédiatement dans le récit homérique le châtiment des amants, enfermés par Vulcain dans un filet.

> Pour nuancer un peu le caractère subversif de ce geste, voir Françoise Gury, 'Mars et Vénus : l'amour dans la dignité : le dossier de la peinture campanienne', in De la théâtralité du corps aux corps des dieux dans l'Antiquité, ed. by Valérie Huet and Florence Gherchanoc (Brest: Centre de Recherche Bretonne et Celtique, 2014), pp. 143-161

Figure 10. Discussion leading to modification of a footnote in word processor.

repertoire of a given language, the `<etym>` element will obviously appear as a child of `<entry>`. It can also be the case that one has to deal with the emergence of a new sense for a given word, in which case `<etym>` should be attached to the corresponding `<sense>` element. Clearly, this distinction is correlated with etymological types, borrowing being more likely to be related to lexical entries whereas metaphors would more often correspond to new senses.

Any change that occurs within a lexicon can be labeled in the `@type` attribute of the `<etym>` element in a TEI dictionary;¹⁶ and the fact that a change occurred within the contemporary lexicon (as opposed to its parent language) is indicated by means of `@xml:lang` on the source form.¹⁷

> Oups! The ticket has been accepted by the council. See https://github.com/TEIC/TEI/issues/1512 We need to change the footnote to say that we made the proposal and that it is about to be implement in a future version of the guidelines

> Please see if my revisions to the footnote make sense.
>
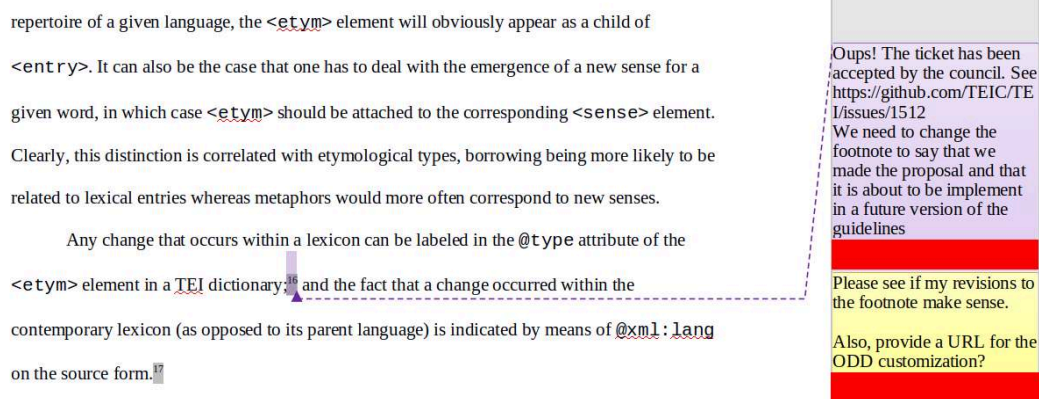> Also, provide a URL for the ODD customization?

Figure 11. Modification of a footnote following a discussion in word processor.

¹⁶ Currently the use of @type in the `<etym>` element is not permitted in the TEI schema as the aforementioned element is not a member of the `att.typed` class. However, a~~The~~ proposal for this use has been ~~submitted in~~accepted by the TEI Council ~~TEI GitHub~~and will be implemented in a future version of the TEI Guidelines; meanwhile, the ticket~~and~~ is available at https://github.com/TEIC/TEI/issues/1512. In our proposal, the `<etym>` element has to be made recursive in order to allow the fine-grained representations we propose here. The corresponding ODD customization, together with reference examples, is~~are~~ available on GitHub.

¹⁷ There may also be cases in which it is unknown whether a given etymological process occurred within the contemporary language or parent system; in such cases the encoder can just use the main language tag for both the diachronic and synchronic portions of the entry as a default (see, for instance, example 11).

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

17

Figure 12. `RedSchema` **with discussion and modification of a footnote.**


```
<p> Any change that occurs within a lexicon can be labeled in the @type attribute of the @etym element in a TEI dictionary;
    <anchor xml:id="Rev2beg"/>
<note resp="#aut" n="16"> Currently the use of @type in the @etym element is not permitted in the TEI schema as the
aforementioned element is not a member of the att.typed class. <choice><sic resp="#aut"> The </sic> <corr resp="#Ed1">However, a
</corr></choice> proposal
<add resp="#Ed1">for this use</add> has been <choice><sic resp="#aut"> submitted in </sic><corr resp="#Ed1"> accepted by </corr>
</choice> the <choice><sic resp="aut"> TEI GitHub </sic> <corr resp="#Ed1"> TEI Council </corr></choice> <add resp="Ed1"> and
will be implemented in a future version of the TEI Guidelines; meanwhile, the ticket </add> <del resp="#Ed1"> and </del> is
available at https://github.com/TEIC/TEI/issues/1512. In our proposal, the @etym element has to be made recursive in order to
allow the fine-grained representations we propose here. The corresponding ODD customization, together with reference examples, <
choice> <sic resp="#aut"> are </sic> <corr resp="#Ed1"> is </corr></choice> available on GitHub.</note>
<anchor xml:id="Rev2end"/>
and the fact that a change occurred within the contemporary lexicon (as opposed to its parent language) is indicated by means of
@xml:lang on the source form.2 </p>
</div>
        <app xml:id="Rev2" from="#Rev2beg" to="#Rev2end"> <note resp="#aut"> Oups! The ticket has been accepted by the council. See
        <ref target="https://github.com/TEIC/TEI/issues/1512"> TEIC/TEI/issues/1512 </ref>
We need to change the footnote to say that we made the proposal and that it is about to be implement in a future version of the
guidelines </note> </app>


        <div type="discussion">
<p resp="#Ed1"> <ptr target="#Rev2"/>Please see if my revisions to the footnote make sense. Also, provide a URL for the ODD
customization??</p>
        </div>
```

Figure 13. `<teiHeader>` **of examples in** figures 7 and 12.


```
1   <?xml version="1.0" encoding="UTF-8"?>
2   <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml" schematypens="
    http://relaxng.org/ns/structure/1.0"?>
3   <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml"
4   schematypens="http://purl.oclc.org/dsdl/schematron"?>
5   <TEI xmlns="http://www.tei-c.org/ns/1.0">
6     <teiHeader>
7       <fileDesc>
8         <titleStmt>
9           <title>                                              /title>
10          <author xml:id="aut">Anonyme</author>
11          <respStmt>
12            <resp xml:id="Ed1">reviewer A</resp>
13            <name>Anonyme</name>
14          </respStmt>
15          <respStmt>
16            <resp xml:id="Ed2">reviewer B</resp>
17            <name>Anonyme</name>
18          </respStmt>
19          <respStmt>
20            <resp xml:id="Ed3">reviewer C</resp>
21            <name>Anonyme</name>
22          </respStmt>
23          <respStmt>
24            <resp xml:id="Ed4">reviewer D</resp>
25            <name>Anonyme</name>
26          </respStmt>
27        </titleStmt>
28        <publicationStmt>
29          <p>          </p>
30        </publicationStmt>
31        <sourceDesc>
32          <p>        </p>
33        </sourceDesc>
34      </fileDesc>
35    </teiHeader>
36    <text>
37      <body>
38
39
```

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

18

Figure 14. Examples in figures 7 and 12.

```
39
40   <div type="texte" resp="#aut">
41
42   <p>A major milestone in the digital dictionary era is the work by Amsler and Tompa (1988), which, together with the unifying
     contributions of Ide and Véronis (1994), led to the <anchor xml:id="Rev1beg"/>earlier<anchor xml:id="Rev1end"/> TEI "Print
     Dictionaries" chapter. Focusing here on their contribution to etymology, example 1 shows how they have introduced a highly
     structured model based on etymons and links implemented as an SGML  DTD.  The underlying model is based upon a graph of etymons
     (@etymon) connected with relations (@rel), forming a more global etymological tree (reflected by the @es, etymological segment,
     element).
43
44   [...]
45
46   Any change that occurs within a lexicon can be labeled in the @type attribute of the @etym element in a TEI dictionary;
47       <anchor xml:id="Rev2beg"/>
48   <note resp="#aut" n="16"> Currently the use of @type in the @etym element is not permitted in the TEI schema as the
     aforementioned element is not a member of the att.typed class. <choice><sic resp="#aut"> The </sic> <corr resp="#Ed1">However, a
     </corr></choice> proposal
49   <add resp="#Ed1">for this use</add> has been <choice><sic resp="#aut"> submitted in </sic><corr resp="#Ed1"> accepted by </corr>
     </choice> the <choice><sic resp="aut"> TEI GitHub </sic> <corr resp="#Ed1"> TEI Council </corr></choice> <add resp="Ed1"> and
     will be implemented in a future version of the TEI Guidelines; meanwhile, the ticket </add> <del resp="#Ed1"> and </del> is
     available at https://github.com/TEIC/TEI/issues/1512. In our proposal, the @etym element has to be made recursive in order to
     allow the fine-grained representations we propose here. The corresponding ODD customization, together with reference examples, <
     choice> <sic resp="#aut"> are </sic> <corr resp="#Ed1"> is </corr></choice> available on GitHub.</note>
50   <anchor xml:id="Rev2end"/>
51   and the fact that a change occurred within the contemporary lexicon (as opposed to its parent language) is indicated by means of
     @xml:lang on the source form.2 </p>
52
53   </div>
54
55   <app xml:id="Rev1" from="#Rev1beg" to="#Rev1end"><note resp="#Ed1">          Is this the right term?</note>
56           <note resp="#Ed2">"earlier" seems fine to me.</note>
57         <note resp="#Ed3">This would be a good place for a specific citation to that earlier chapter.</note></app>
58
59   <app xml:id="Rev2" from="#Rev2beg" to="#Rev2end"> <note resp="#aut"> Oups! The ticket has been accepted by the council. See <ref
     target="https://github.com/TEIC/TEI/issues/1512"> TEIC/TEI/issues/1512 </ref>
60   We need to change the footnote to say that we made the proposal and that it is about to be implement in a future version of the
     guidelines </note> </app>
61
62
63   <div type="discussion">
64   <p resp="#aut"> <ptr target="#Rev1"/>"early"?</p>
65   <p resp="#Ed1"> <ptr target="#Rev2"/>Please see if my revisions to the footnote make sense. Also, provide a URL for the ODD
     customization??</p>
66           </div>
67
68   </body>
69   </text>
70   </TEI>
71
```

47   To generate a clean text, in the end, the last modification would be considered as final, which means that the last copyeditor should validate with their @resp attribute the earlier modification layers according to the final editorial decision.

48   The main advantage of such an encoding system is that it sums up in one file all the editorial evolution of a text, from submission to publication, displaying precisely its evolution and the contribution of each one of those who were involved in this process. It is, therefore, a question of including within the chain that structures and edits content the part corresponding to the evaluation and formatting, and thus making it visible and shareable. In this way, the data are made open to the point of the preparation of the data themselves. It is worth noting that recent efforts made by scientific publishers to automate workflows have focused on publishing content but did not include the preparation of such content.[36]

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward
a Standardized Workflow

19

49   All of this is summed up not only in one file but also in one unique file format, guaranteeing conversion without information loss. It allows splitting the editorial work on form and content, giving the editor the final hand on the last version of the text. And finally, it should be conceived as a fairly minimal combination of tagsets, meaning that these schemas should be easy to share with other scholars and journals.

50   It would have, on the downside, the inconvenience that comes with its advantages: being multilayered, such a document might quickly become complex. Transformation scenarios filtering specific tagsets to gain readability will be made necessary: both readability for the human eye and information extraction for digital tools would rely on the development of such transformation scenarios. But all in all, the development of such schemas and transformation scenarios seems in the realm of the doable considering what the TEI has been able to develop over the last three decades.

4. Structural Workflow Improvements beyond the TEI

51   The main reason why corrections are implemented, and the correction process hidden from the public eye, is a need for control. Editors and authors want to have control over each modification in the text, be it punctuation, bibliography formatting, or a sentence that seems a bit difficult to understand. This is all the more important in the case of papers written by nonnative speakers. Those need close linguistic attention to reach the same level of readability as those written by native speakers.

52   But looking at it more closely, the whole workflow inherent to this control of modifications and corrections is based on reputation control. It is based on editorial "black boxes" that keep articles away from the public eye as long as they are not "perfect" or "finished." It is based on the idea that work in progress can damage reputation. This is exactly what started to change with the publishing of digital scholarly editions online. The fact that it is possible to update a digital scholarly edition suggests that we could at least imagine that such plasticity can be envisioned for journal articles too. Making it possible to work with sources that are not fixed in time is one of the greatest instances of intellectual progress not only made possible but actually realized by the TEI. Taking it one step further for scholarly journals is a fascinating intellectual challenge in terms of data dissemination quality assurance.[37] It is also the logical next step in terms of data empowerment.

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

20

53   This could easily be encouraged by two (infra)structural lines of action. The first one is to further foster pre-print publications (JTEI encourages such publications for papers submitted to it). The second one consists in improving the academic recognition of editorial expertise. There are already many experts in TEI working as research engineers or editorial assistants. They are in general working for specific journals: for example, for an editor or publisher. This expertise should be better recognized than it is now, and valued more explicitly in advertisements for editorial jobs. The TEI community should encourage this expertise to be better represented, for instance in the special interest groups (SIGs) or through the awarding of prizes. The academic publishing market has been intensely professionalized in terms of digital competence for several years, particularly through the emergence of networks such as the Medici network,[38] which has helped in the continued training of editors. The profession is now ready to integrate such a workflow and support its development.

54   If we were to alter the workflow in such a way that the versioning occurring alongside reviewing and copyediting was easily manageable, it would mean that successive updated versions could be published online in preprint archives until the journal publication appears. For instance, it would be possible to filter through customized transformations what would be suitable for predefined publication steps up to the final version. It would be possible to tailor the review process to make it completely open, open for `EdSchema`, open for `RedSchema`, or not open at all, according to the author's wish and/or the journal's policy. The same goes for copyediting.

55   In this situation where the dialogue between authors, reviewers, and editors can be made transparent within the text, scholars would be in a different position than that of accepting or refusing a correction suggestion. The dynamics of the relationships between those involved in the process of generating the text could benefit from this change. The TEI community could be actively involved in providing schematrons, stylesheets, and publishing environments for journals that would allow researchers to access TEI documents directly, be it their own for additional editorial work or others' for queries.

56   One relatively easy way to implement such a workflow consists in building overlay journals like those hosted on the Episciences platform,[39] which only provides an additional review layer on top of preprint publications. The interface allows a journal to set up a review process and link to the evaluated and selected preprints to make up a journal issue. While this greatly simplifies the

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward a Standardized Workflow

21

review process, especially because of the easy-to-adjust interface, it does not really address the online presentation of publications in a reader-friendly way, for which other solutions have to be implemented and which would require an additional investment.

5. Conclusion

57 Workflows are complicated because, in theory, they address all the needs of the editorial process. In terms of scholarly publishing workflows, there is some development work to do before we will be able to disseminate an encompassing TEI schema for authoring, reviewing, and publishing. The encoding we have proposed here could serve as a basis; potential users are already working with a TEI-based publishing format, and the TEI community has a great deal of expertise to build upon. Looking at the many journals that are already working with a TEI-based publishing format, there is one major thing required to take it to the next step: good documentation everywhere.

58 The TEI community is in a position to impact access to knowledge for future generations—not only through our digital editions of manuscripts, but also through the way we disseminate all the information we gather from working with manuscripts and with digital editions. Fostering TEI-based scholarly publications is worthwhile: first, it has been done at different scales for publishing and is working; second, our knowledge of it allows us to reflect on the specificities of the different disciplines we come from; and third, it is up to us to initiate this change because nobody else can do it the way the TEI community can. There is no reason why the coming generations should be plagued by time- and money-consuming requirements from publishing houses now that there is a real political awareness of the need for solid archiving formats and now that the TEI has established itself as a standard in so many fields.

BIBLIOGRAPHY

Bauman, Syd. 2011. "Interchange vs. Interoperability." In *Proceedings of Balisage: The Markup Conference 2011.* Balisage Series on Markup Technologies 7. doi:10.4242/BalisageVol7.Bauman01.

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward
a Standardized Workflow

22

Fitzpatrick, Kathleen, and Avi Santo. 2012. *Open Review: A Study of Contexts and Practices.* Andrew W. Mellon Foundation White Paper. Final version, December. Produced by MediaCommons and NYU Press for the Andrew W. Mellon Foundation. http://mcpress.media-commons.org/open-review/; https://mellon.orgmedia/filer_public/20/ff/20ff03e0-17b0-465b-ae82-1ed7c8cef362/mediacommons-open-review-white-paper-final.pdf.

Girard, Chloé. 2017. "Les mécanismes de centralisation des données de la recherche: Étendre l'accès libre à l'hébergement libre" [Mechanisms of Centralization of Research Data: Extending Open Access to Open Hosting]. *Revue française des sciences de l'information et de la communication*, no. 11. doi:10.4000/rfsic.3255.

Holmes, Martin, and Laurent Romary. 2011. "Encoding Models for Scholarly Literature: Does the TEI Have a Word to Say?" In *E-Publishing and Digital Libraries: Legal and Organizational Issues*, edited by Ioannis Iglezakis, Tatiana-Eleni Synodinou, and Sarantos Kapidakis, 88–110. Hershey, PA: IGI Global. doi:10.4018/978-1-61692-834-6.ch005. Authors' version available at https://hal.archives-ouvertes.fr/hal-00390966v2 .

Homenda, Nicholas, and Shayna Pekala. 2016. "Migrate, Publish, Repeat: TEI Journals in the Open Journal Systems Platform." *Journal of the Text Encoding Initiative* 10. http://journals.openedition.org/jtei/1746; doi:10.4000/jtei.1746.

Kosmopoulos, Christine, and Denise Pumain. 2008. "Révolution numérique et évaluation bibliométrique dans les sciences humaines et sociales." *Revue européenne des sciences sociales*, no. XLVI-141: 73–86. doi:10.4000/ress.151.

Langlais, Pierre-Carl. 2016. "Étude critique des nouveaux modes 'd'éditorialisation' de revues scientifiques en accès-ouvert." Research report, Bibliothèque Scientifique Numérique (Digital Scientific Library). https://hal.archives-ouvertes.fr/hal-01388556.

Pytlik Zillig, Brian L. 2009. "TEI Analytics: Converting Documents into a TEI Format for Cross-Collection Text Analysis." *Literary and Linguistic Computing* 24 (2): 187–92. doi:0.1093/llc/fqp005.

Thoden, Klaus. 2019. "Modeling Scholarly Publications for Sustainable Workflows." *ELPUB 2019. 23rd edition of the International Conference on Electronic Publishing.* June 2019, Marseille, France. hal-02143240, v. 1. https://elpub.episciences.org/5560/; doi:10.4000/proceedings.elpub.2019.2.

TEI Models for the Publication of Social Sciences and Humanities Journals: Opportunities, Challenges, and First Steps Toward
a Standardized Workflow

23

Unsworth, John. 2011. "Computational Work with Very Large Text Collections: Interoperability, Sustainability, and the TEI." *Journal of the Text Encoding Initiative* 1. https://journals.openedition.org/jtei/215; doi:10.4000/jtei.215.

NOTES

1 As explained on the TEIWiki page "JTEI Authoring Tools," last edited May 14, 2015, https://wiki.tei-c.org/index.php?title=JTEI_Authoring_Tools.

2 See Anne Baillot, "Scholarly journals publishing in TEI," TEI-L archives, August 22, 2019, and replies, https://listserv.brown.edu/cgi-bin/wa?A1=ind1908&L=TEI-L#9.

3 If the topic gains momentum, it would be thinkable to reactivate the currently dormant TEI Special Interest Group (SIG) on Scholarly Publishing and manage to keep the information on the SIG's wiki up to date.

4 Accessed February 15, 2021, https://www.openedition.org/.

5 The schema documentation is found on L'Édition Électronique Ouverte, one of the Hypotheses blogs hosted by the platform: see Jean-François Rivière, "Schema XML TEI OpenEdition," last updated September 15, 2014, https://leo.hypotheses.org/10717.

6 Current figures available through the search interface: https://search.openedition.org/results?s=&pf=OJ.

7 Current figures available through the search interface: https://search.openedition.org/results?s=&pf=OB.

8 Information from Pierre Mounier, private communication.

9 Accessed February 15, 2021, https://hal.archives-ouvertes.fr/.

10 Accessed February 15, 2021, https://anr.fr/fr/lanr-et-la-recherche/engagements-et-valeurs/la-science-ouverte/.

11 Agnes Magron, "L'Hcéres choisit la plateforme HAL pour l'archivage et la diffusion des rapports d'évaluation," Centre pour la Communication Scientifique Directe, April 9, 2019, https://www.ccsd.cnrs.fr/2019/04/hceres-choisit-la-plateforme-hal-pour-archivage-et-la-diffusion-des-rapports-evaluation/.

12   Individual researchers are to use HAL imports in the CRAC system (accessed [march 2021], https://www.science-ouverte.cnrs.fr/les-actions-du-cnrs/) from 2019 on and in the RIBAC system for SSH scholars (accessed February 15, 2021, https://www.ribac-shs.cnrs.fr/) from 2020 on.

13   A rather vigorous discussion has been going on since the early 2000s on this subject. It concerns on the one hand the centralizing approach of HAL, and on the other hand the progressive use of the open archive for bibliometric purposes, in order to produce objective indicators intended to direct the public policy of support for research. For a record of the first branch of the discussion, see the debate between Stevan Harnad and Franck Laloë on the American Scientist Open Access Forum during 2006 (Franck Laloe, "Re: France's HAL, OAI interoperability, and Central vs Institutional Repositories," October 3, 2006, https://web.archive.org/web/20120403072426/http://listserver.sigmaxi.org/sc/wa.exe?A2=ind06&L=american-scientist-open-access-forum&D=1&O=D&F=l&S=&P=75655); for a record of the second branch, see the conversation that took place on the French open access list on June 25, 2018, entitled "Fichiers audio de: 'HAL et les autres archives ouvertes: libre accès aux savoirs, centralisation des données et quantifications néo-managériales ou marketing académique?' séminaire PDS/EHESS, 14/06/2018 18h–21h" (https://groupes.renater.fr/sympa/arc/accesouvert/2018-06/msg00002.html). The problem is very well addressed by Girard (2017).

14   Accessed February 23, 2021, https://www.istex.fr/.

15   See Pascale Viot, "American Chemical Society' ou ACS désormais disponible sur la plateforme ISTEX," ISTEX blog, August 2, 2019, https://blog.istex.fr/american-chemical-society-ou-acs-desormais-disponible-sur-la-plateforme-istex/.

16   *Nationale Forschungsdateninfrastruktur (NFDI), Deutsche Forschungsgemeinschaft*, last updated January 11, 2021, https://www.dfg.de/foerderung/programme/nfdi/.

17   DARIAH (Digital Research Infrastructure for the Arts and Humanities), accessed February 24, 2021, https://www.dariah.eu/.

18   CLARIN—European Research Infrastructure for Language Resources and Technology, accessed February 24, 2021, https://www.clarin.eu/.

19   On ERICs in general, see "European Research Infrastructure Consortium (ERIC)," European Commission, accessed February 24, 2021, https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures/eric_en.

20   Accessed March 21, 2021, http://www.zfdg.de/; 90 articles available as of March 21, 2021.

21   RIDE: A Review Journal for Digital Editions and Resources, accessed February 24, 2021, https://ride.i-d-e.de/.

22   "Über Uns," Institut für Dokumentologie und Editorik, accessed February 24, 2021, https://www.i-d-e.de/uber-uns-about-us/.

23   See the RIDE GitHub repository, accessed February 24, 2021, https://github.com/i-d-e/ride. There is also an evaluation of all questionnaires in TEI XML available for further usage (https://github.com/i-d-e/ride/tree/questionnaire_data). The schema documentation was being revised at the time of writing this article; it is based on the jTEI Article schema and only deviates for specific aspects related to the nature of the publication.

24   "Jahrbuch," Deutsche Gesellschaft für Erziehungswissenschaft, accessed February 24, 2021, https://www.dgfe.de/sektionen-kommissionen-ag/sektion-1-historische-bildungsforschung/jahrbuch

25   Accessed February 14, 2021, http://ojs.inz.si/pnz.

26   See the journal's home page, accessed February 11, 2021, https://hugoye.bethmardutho.org/index.html.

27   Accessed February 11, 2021, http://romantic-circles.org/about/about.html.

28   Tony Ross-Hellauer, Benedikt Fecher, Kathleen Shearer, and Eloy Rodrigues, "Pubfair: A Distributed Framework for Open Publishing Services," v. 2, November 27, 2019, https://www.coar-repositories.org/news-updates/pubfair-version-2-now-available/.

29   On the TEI's proposed compromise between precision and ability to share, see Bauman (2011). On the challenge posed to the TEI by the notion of interoperability and on the need to build a general ecosystem conducive to achieving the full ability to share primary or secondary documents, see Unsworth (2011). Automatic conversion of SSH documents into a common form of markup is currently being explored by the MONK project, as described in Pytlik Zillig (2009). The Abbot transformation tool developed by the members of the MONK project opens promising perspectives for the harvesting of very important XML corpuses, and thus contributes to solving the issue of TEI document searchability, which is still important today.

30   Among these attempts we can list the traditional OJS that is the basis of many small and medium-sized editorial endeavors, but also major journal hosting initiatives such as Érudit (accessed February 25, 2021, https://www.erudit.org/en/); other examples are described above in section 1.

31   See "FAIR Principles," GO FAIR, accessed February 25, 2021, https://www.go-fair.org/fair-principles/.

32   See also Thoden (2019) on the strategy of converting the workflow and document basis from a proprietary format to a fully standards-compliant system in the context of a publishing platform.

33   See the journal's OpenEdition page, accessed February 25, 2021, https://journals.openedition.org/philosant/.

34   We would like to thank the authors, reviewers, and copyeditors who allowed us to use for this purpose data that are usually not made public but considered part of the "black box" of academic publishing.

35   It should be possible to add a bibliographical entry using the `<resp>` to indicate that it comes from a specific reviewer or a copyeditor.

36   See, for instance, the METOPES XML editorial workflow, which is based on the principle of single-source publishing but leaves aside the preparation of content itself: "Environnement Métopes," pôle Document numérique, Maison de la recherche en sciences humaines (CNRS / Université de Caen Normandie), accessed February 25, 2021, http://www.unicaen.fr/recherche/mrsh/document_numerique/outils/metopes, which is not making its schemas openly available at this stage.

37   Easy access to all the data from the editing process in a structured format could be a decisive step in promoting open peer-review practices. Langlais (2016) thus questions the relationship between the success of recent attempts at open review and the tools used, and suggests that new devices in this direction remain rare and experimental (26–27). We can also refer to Fitzpatrick and Santo, who propose an interesting specification of a tool or an ecosystem of tools favorable to open peer review; the notion of aggregation ("An open review system must be able to gather and display the complete contributions of an individual user") (2012, 24) seems particularly interesting to us and convergent with what we propose.

38 "Médici (Métiers de l'ÉDition sCIentifique publique)," accessed February 25, 2021, http://medici.in2p3.fr/?lang=fr.

39 Accessed February 25, 2021, https://www.episciences.org/.

AUTHORS

ANNE BAILLOT

Anne Baillot is a full professor in German studies at Le Mans Université. She is the editor of *Letters and Texts. Intellectual Berlin around 1800* and former managing editor of the *Journal of the Text Encoding Initiative.*

JULIE GIOVACCHINI

Julie Giovacchini is a research engineer in Classics at French CNRS (Centre Jean Pépin UMR8230 CNRS-ENS-PSL). She is co-editor of *Philosophie antique*, a journal dedicated to the history of ancient philosophy.