Journal of the Text Encoding Initiative

Issue 14 | 2021
Selected Papers from the 2019 TEI Conference

Analyzing and Visualizing Uncertain Knowledge: The Use of TEI Annotations in the PROVIDEDH Open Science Platform

Michał Kozak, Alejandro Rodríguez, Alejandro Benito-Santos, Roberto Therón, Michelle Doran, Amelie Dorn, Jennifer Edmond, Cezary Mazurek and Eveline Wandl-Vogt

Electronic version
URL: https://journals.openedition.org/jtei/4239
DOI: 10.4000/jtei.4239
ISSN: 2162-5603

Publisher
TEI Consortium

Electronic reference
Michał Kozak, Alejandro Rodríguez, Alejandro Benito-Santos, Roberto Therón, Michelle Doran, Amelie Dorn, Jennifer Edmond, Cezary Mazurek and Eveline Wandl-Vogt, “Analyzing and Visualizing Uncertain Knowledge: The Use of TEI Annotations in the PROVIDEDH Open Science Platform”, *Journal of the Text Encoding Initiative* [Online], Issue 14 | 2021, Online since 13 September 2022, connection on 15 September 2022. URL: http://journals.openedition.org/jtei/4239 ; DOI: https://doi.org/10.4000/jtei.4239

Analyzing and Visualizing Uncertain Knowledge: The Use of TEI Annotations in the PROVIDEDH Open Science Platform

Michał Kozak, Alejandro Rodríguez, Alejandro Benito-Santos, Roberto Therón, Michelle Doran, Amelie Dorn, Jennifer Edmond, Cezary Mazurek, and Eveline Wandl-Vogt

ABSTRACT

The underlying uncertainty in digital humanities research data affects decision-making and persists during a project's lifecycle. This uncertainty is inevitable since most empirical claims cannot be assessed against an absolute truth (Drucker 2011; Binder et al. 2014). This situation has been previously recognized together with the need to report the degrees of uncertainty that accompany such claims (Blau 2011). Although TEI makes it possible to annotate text with notions of certainty or precision, examples of actual projects taking advantage of this are scarce. There are many possible explanations for uncertainty's lack of visibility in computationally supported humanities research; among them, the need for tools specifically designed to address the goal of defining and managing uncertainty stands out. Thus, efforts to provide technical support for humanities research should focus on managing and making uncertainty more transparent, rather than removing it. Another challenge is the fact that there is no agreement on a generic taxonomy for the different types of uncertainty that researchers may face. Various researchers across disciplines, working on varying projects and data sets, can use different categories to classify the uncertainties present in a particular case.

In this paper, we introduce a collaborative platform for collective annotation of TEI data sets. We briefly present the flexible taxonomy of uncertainty used in the platform and describe two data sets used for its testing. Then we describe use cases of annotations available on the platform, and how they translate into TEI annotations. Creating and interpreting annotations with and without uncertainty should now be easier, especially for researchers who do not know TEI markup.

INDEX

Keywords: uncertainty, annotation, visualization, collaborative annotating

ACKNOWLEDGEMENTS

This work has received funding within the CHIST-ERA Programme under the following national grant agreements: 1. Spain, PCIN-2017-064, MINECO; 2. Poland, 2017/25/Z/ST6/03045, National Science Centre; 3. Austria, FWF – I 3441 – N33, FWF Austrian Science Fund; and 4. Ireland, Irish Research Council.

1. Introduction

1 The management and display of uncertainty in computational processes has recently gained interest within the scientific community. As the volumes of data and algorithms at play grow in complexity, many scholars are becoming aware of the importance of communicating uncertainty as a means to enhance the explainability and accountability of computational processes in a wide range of domains (Hullman 2020). As opposed to previous approaches, which mainly focused on the systematic elimination of uncertainty by applying different statistical constructs, modern

solutions tend to provide faithful representations of the underlying statistical models with the aim of enabling better comprehension by end users, ultimately allowing them to make more informed decisions based on the reality of the observed data.

2 In this regard, the digital humanities (DH) community has approached the problem from various angles, such as the development of data quality metrics for digital cultural collections (Windhager et al. 2019), the quantification and categorization of uncertainty in DH analysis processes (Martín-Rodilla and González-Pérez 2019; Rocha Souza et al. 2019), the improvement of existing strategies for the development of digital research tools and environments for history (Edmond 2019), and from the perspective of confidence (Franke et al. 2019). In the course of this work, manual and automatic annotation have emerged as one of the junctions in DH workflows where uncertainty can be introduced. Despite its importance, the consideration of uncertainty in the annotation task has not been thoroughly addressed in the literature. Thus, this situation has brought the attention of interdisciplinary teams of researchers seeking to understand the implications of bringing uncertainty to the surface of DH projects. This is the case with PROVIDEDH (PROgressive VIsual DEcision-Making in Digital Humanities),[1] an EU-funded four-year project that aims to provide DH scholars with an online space to consider uncertainty related to the completeness and evolution of digital research objects in order to enhance the development and corroboration of hypotheses in a wide range of use cases. One of the project's outcomes is the development of a collaborative platform centered around the collaborative, uncertainty-aware annotation of TEI texts.[2]

3 Fortunately, the identification and tracing of certain types of uncertainty in the life cycle of a textual corpus can be partially achieved with the use of TEI tags (TEI Consortium 2020, ch. 21, Certainty, Precision, and Responsibility).[3] However, these methods are not commonly used. The motivation of this paper is to enable a broader landscape of use cases in which these tags can be utilized by exposing them to the end user in a bespoke user interface. In particular, the platform facilitates a use of the `<certainty>` tag that gives the possibility of annotating the degree of uncertainty with an attribute `@cert` that allows the assignment of broad categorical values ("high" / "medium" / "low" / "unknown") or `@degree` to allow for more granular numerical values. The platform provides a user-friendly interface for annotating entities and uncertainties related to them, so it mainly exploits `@cert` , which is intended more for humans. However, it also allows

users to upload already annotated data sets in which @degree attributes are filled in algorithmically (because it is likely impossible for humans to ascertain a difference between, for instance, 0.67 and 0.68 degrees of confidence).

4 The ultimate goal of this implementation of <certainty> and its attributes is to provide, for teams which collaboratively edit texts, an instrument to explicitly document their assumptions and conclusions regarding uncertain passages in their shared text. Alternatively, these tags can assist in making TEI-encoded texts reusable beyond the FAIR (findable, accessible, interoperable, and reusable) criteria for data. This deployment of <certainty> can instead make the data more epistemically available, with the provenance or prior interpretive effort that has been applied to it, integrated into the editing process that shaped it.

2. Uncertainty Taxonomy

5 One of the first tasks of the PROVIDEDH project was to describe and to define the various sources of uncertainty in text-based DH research objects, with the aim of compiling a taxonomy of uncertainty for visualization and representation of uncertainty in DH. This taxonomy is now stable, though it remains under development as a living document. In its current iteration, it is divided into two main categories: “user-recognized” and “machine-generated” uncertainty. The former describes the manner in which users perceive and categorize uncertainty in digital texts and is related to previous work in visualization and the humanities (Fisher 1999; Simon, Weber, and Sallak 2018). The latter concerns cases when the annotation (including about uncertainty) is done by an algorithm. For a complete explanation of the motivations behind the proposed taxonomy and its categories we refer the reader to Therón Sánchez et al. (2019). Although initially the taxonomy proposed a closed, fixed set of categories, we are currently departing from this approach because of a series of recent findings resulting from interactions with real users. During these preliminary evaluations, we discovered, for example, that the optimal naming of the categories fluctuates depending on the project’s specific humanities content and its end user’s academic background (e.g., certain users prefer to employ “imprecision” while others refer to the same category as “vagueness”). Furthermore, recent findings are pointing us to think that the number of human-assigned categories (high, medium, low, and unknown) could, in some cases, not be exactly four but rather vary slightly around this figure. Finally, and in order to augment the expressive capabilities

of the taxonomy, we are also supporting a flexible (or fuzzy) implementation that allows users to combine two or more categories in a single statement. In summary, although the capture of user-perceived uncertainty in digital interfaces is a very sensitive matter that necessitates implementing important validation strategies, in this paper we report on preliminary strategies at the data format level that we know will be required for our platform to meet the identified user needs.

6 Finally, the taxonomy and platform are currently being tested in two separate scenarios: (1) free annotation of texts, and (2) the normalization/unification of entities, which involves the assertion by a human actor, with some degree of certainty, that two or more entities are indeed the same. In our tests, we employed two historical, TEI-encoded data sets that are described in the next section.

7 We are aware that marking uncertainty levels and categories is a functionality unlikely to be used by many individual scholars working within current paradigms, and indeed we would imagine many potential users might not take advantage of the affordances of such an option set. Most scholars may just want to mark identifications or readings as uncertain but will not want to qualify this uncertainty further. But our approach emerged within a project and platform context where the convergence of a number of different emerging research paradigms demonstrated a strong case for the inclusion of such values. First, they can play a role in the support of collaborative and indeed distributed scholarship. The characterization of something as uncertain always implies a contextual richness that markup is ill-equipped to capture, in particular as uncertainty is ultimately an epistemic state, and therefore resides in the individual consciousness of the scholar. Where that richer tacit information may not be directly available because, for example, two scholars are working together without regular direct communication, markers of the degree or categories of uncertainty can take on a significant signaling function without needing to capture too much detail. Second, the decision to include these tags emerged within a project that was testing the potential of progressive visualization as a research tool for historians. Considering and assigning degrees and categories of uncertainty may not seem to be a particularly valuable exercise in and of itself, unless or until a tool such as a visualization can allow these attributes of uncertainty to be viewed in a comparative framework and used to weigh alternative explanations, distinguish norms from outliers, and prioritize future efforts.

3. Data Sets Description

3.1 1641 Depositions

8 The 1641 Depositions (Trinity College Dublin, MSS 809–841)[4] are witness testimonies mainly by Protestants, but also by some Catholics, from all social backgrounds, concerning their experiences of the 1641 Irish rebellion. The testimonies document the loss of goods, military activity, and alleged crimes committed by the Irish insurgents, including assault, stripping, imprisonment, and murder. This body of material is unparalleled anywhere in early modern Europe, and provides a unique source of information for the causes and events surrounding the 1641 rebellion and for the social, economic, cultural, religious, and political history of seventeenth-century Ireland, England, and Scotland.

3.2 Historical Recipes

9 A second case study used to test our annotator was a data set of historical recipes of the Baroque era from the former area of Austria and beyond.[5] The recipe collection was previously established and pertained to the citizen science project Salzburg zu Tisch,[6] carried out and led by our cooperation partners, staff members of the Center for Gastrosophie,[7] the History Department at the University of Salzburg, Austria. The Gastrosophie recipe collection counts around ten thousand historical recipes from different cookbooks and different authors, with the majority being in nonstandard German and a small number in other languages such as English. During this previous project, the recipes were digitized, entered into a WordPress instance, and annotated with the help of interested citizens. In this project, they the recipes have been converted to TEI[8] and collectively further annotated (see figure 2).

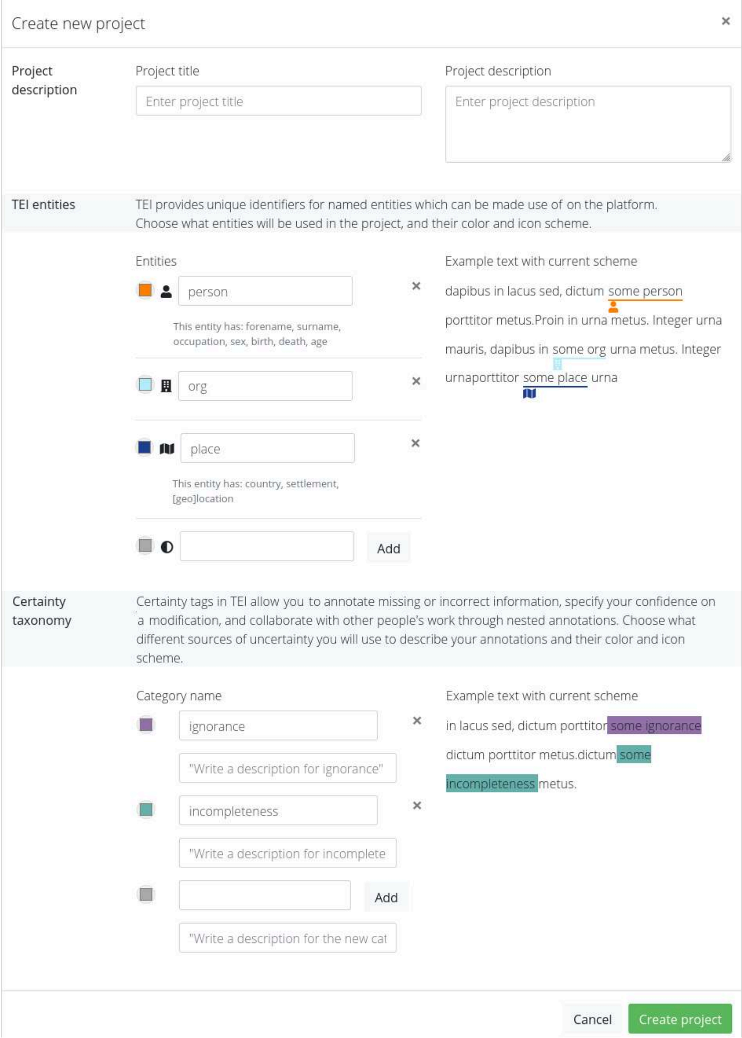
4. Collaborative Platform

10 Given the broad spectrum of named entities that a project may require, based on the data set being used and the variation in what different users consider to be a meaningful certainty-category naming, we identified a need to allow the user to specify the entities and uncertainty categories

within the project being created in the platform. Such a specification determines what entities and their properties will be available and how users will assess the correctness of the documents. This must be set up before working with the TEI documents, during project creation.

11 For this purpose, we provide an interactive interface (see figure 1) that abstracts how the corresponding TEI formatting will be done, and which allows the user to specify both the entities and the taxonomy as well as the colors and icons that will be used across the application to encode this information.

Figure 1. Screenshot of project creation settings.



Create new project ×

Project description

Project title: Enter project title

Project description: Enter project description

TEI entities

TEI provides unique identifiers for named entities which can be made use of on the platform.
Choose what entities will be used in the project, and their color and icon scheme.

Entities

- person ×
  This entity has: forename, surname, occupation, sex, birth, death, age
- org ×
- place ×
  This entity has: country, settlement, [geo]location
- Add

Example text with current scheme

dapibus in lacus sed, dictum some person porttitor metus.Proin in urna metus. Integer urna mauris, dapibus in some org urna metus. Integer urnaporttitor some place urna

Certainty taxonomy

Certainty tags in TEI allow you to annotate missing or incorrect information, specify your confidence on a modification, and collaborate with other people's work through nested annotations. Choose what different sources of uncertainty you will use to describe your annotations and their color and icon scheme.

Category name

- ignorance ×
  "Write a description for ignorance"
- incompleteness ×
  "Write a description for incomplete
- Add
  "Write a description for the new cat

Example text with current scheme

in lacus sed, dictum porttitor some ignorance dictum porttitor metus.dictum some incompleteness metus.

Cancel Create project

12 Working with existing TEI entities (such as "person," "place" or "event") and user-defined ones (such as "ingredient" or "utensil") is tightly integrated into the proposed system, therefore making the whole process effortless to the user, which in turn makes the platform usable in a broad range of different DH research fields.

13 The configuration of the project requires handling the definition of new entities and the creation of a taxonomy. The first is accomplished by using the TEI `<object>` element with its `@type` attribute. Custom certainty categories are handled by creating a project-specific taxonomy which is then referenced in annotations. In addition to the name, optional descriptions can be added for each category to help users understand how uncertainty is being conceived and used in their project. Such information results in the creation of a declaration like the following (which varies based on the project settings):

Example 1. Uncertainty taxonomy generated from default settings.

```
<taxonomy>
 <category>
  <catDesc>User recognized uncertainty</catDesc>
  <category xml:id="ignorance"><catDesc>Ignorance</catDesc></category>
  <category xml:id="credibility"><catDesc>Credibility</catDesc></category>
  <category xml:id="imprecision"><catDesc>Imprecision</catDesc></category>
  <category xml:id="incompleteness"><catDesc>Incompleteness</catDesc></category>
 </category>
 <category>
  <catDesc>Machine generated uncertainty</catDesc>
  <category xml:id="algorithmic"><catDesc>Algorithmic</catDesc></category>
 </category>
</taxonomy>
```

14 The uncertainty taxonomy generated from project settings is available at `https://providedh.ehum.psnc.pl/api/projects/{project-id}/taxonomy/` on the platform and, as we have mentioned, it is referenced in uncertainty annotations. Other settings (such as colors or icons) are stored in the internal platform database. Storing them in TEI files is not necessary as they are only used for visualizations inside the platform.

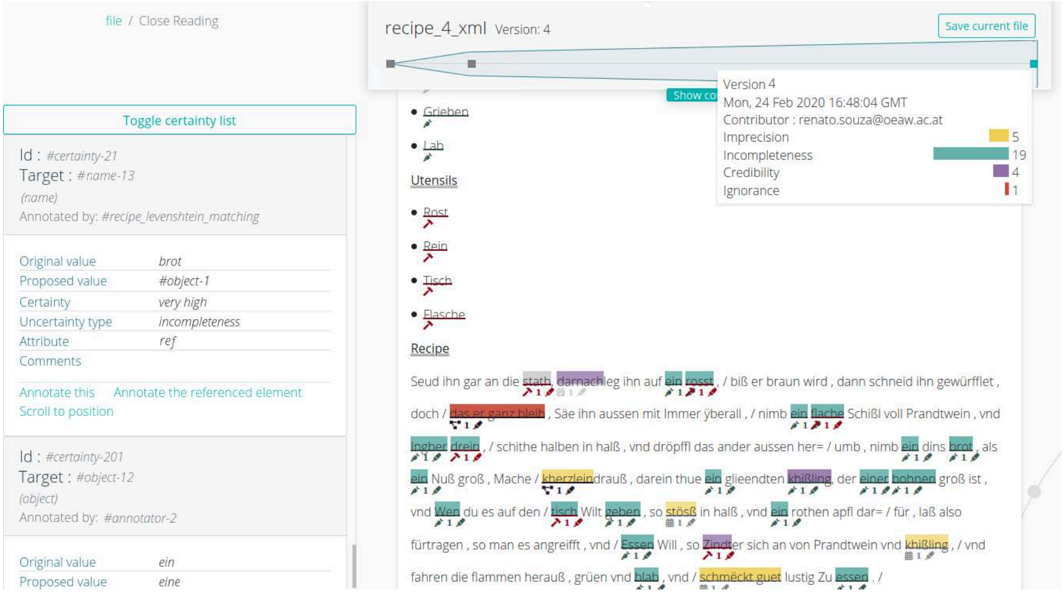
15 After creating a project, the creator can assign other users to the project and give them read or write permissions. The creator starts working on a project by uploading a data set to the platform. The platform supports files in UTF8-encoded TEI P5 format; if a user prefers to start with plain text or use another encoding, the files will be converted to UTF8-encoded TEI P5 during the upload. If TEI files already contain annotated entities, the platform will assign them `@xml:id` so that they can be referred to in further work. Files changed in this way are stored in the platform as a second version of the uploaded files, so the user can track the history of files. In addition, existing annotations and entities are entered into the database, making them easier to manage.

5. Uncertainty Annotator

16 In the aforementioned data sets, we identified a large number of cases where uncertainty can be annotated. This implies that there are a large number of users with some amount of prior knowledge of XML editing and the TEI format, and these uses cases provide a good opportunity to make use of interactive interfaces to ease working with data sets and collaborating within the documents.

17 We, therefore, designed and integrated into the platform an annotator tool that allows for both the reading of TEI documents in a user-friendly manner and the annotating them using simple text-selection interactions. This interface strongly emphasizes collaborative editing, making use of visual encodings to easily distinguish the authorship of annotations at first glance.

Figure 2. Screenshot of the Annotator. In the right pane, the text is displayed with tags colored according to the type of uncertainty annotated by the users. Changes in the uncertainty tags over time are represented in the chart at the top right. In the left pane, the details for each annotation are shown, including the original text to which the annotation refers, the uncertainty degree and type chosen by the author, and other useful information.



Id : #certainty-21
Target : #name-13
(name)
Annotated by: #recipe_levenshtein_matching

Original value	brot
Proposed value	#object-1
Certainty	very high
Uncertainty type	incompleteness
Attribute	ref
Comments	

Id : #certainty-201
Target : #object-12
(object)
Annotated by: #annotator-2

Original value	ein
Proposed value	eine

Version 4
Mon, 24 Feb 2020 16:48:04 GMT
Contributor : renato.souza@oeaw.ac.at

Imprecision	5
Incompleteness	19
Credibility	4
Ignorance	1

- Grieben
- Lab

Utensils

- Rost
- Rein
- Tisch
- Flasche

Recipe

Seud ihn gar an die stath, darnachleg ihn auf ein rosst , / biß er braun wird , dann schneid ihn gewürfflet , doch / das er ganz bleib , Säe ihn aussen mit Immer ÿberall , / nimb ein flache Schißl voll Prandtwein , vnd Ingber drein , / schithe halben in halß , vnd dröpffl das ander aussen her= / umb , nimb ein dins brot , als ein Nuß groß , Mache / kherzleindrauß , darein thue ein glieendten khißling, der einer bohnen groß ist , vnd Wen du es auf den / tisch Wilt geben , so stöß in halß , vnd ein rothen apfl dar= / für , laß also fürtragen , so man es angreifft , vnd / Essen Will , so Zindter sich an von Prandtwein vnd khißling , / vnd fahren die flammen herauß , grüen vnd blab , vnd / schmëckt guet lustig Zu essen . /

18 This collaboration is possible thanks to the use of the TEI format, enabling precise targeting of the source of uncertainty not only regarding the tag name, attributes, or values for entities but also for other people's annotations.

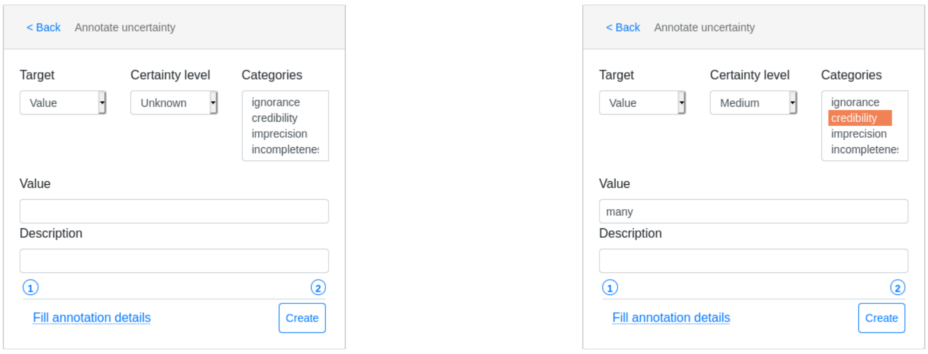
19 The Annotator adds all the changes related to uncertainty to the TEI header. Each user can annotate the same pieces of text in their own way; if we put different users' tags in the same places in the text, it would make the document quite difficult to read. We decided to place all the uncertainty annotations in `<certainty>` elements in the header and only refer to them from the text. We further decided not to use other TEI options for tagging uncertain alternative texts, like `<choice>`, in order to manage uncertainty annotations in a coherent way within the platform. We also considered storing the uncertainty annotations in a `<standOff>` element, which we use to store entities occurring in the text, but ultimately, we decided to store them together with the annotators in the `<profileDesc>` element, as shown in example 2.

20 Next, we describe in depth the use cases for annotation that we identified, and how they are managed within the platform.

5.1 Use Cases A and B: Annotating Some Text as Uncertain with an Optional Alternative Value

21 First, let’s present a modal window that appears after the user selects a text fragment and chooses the “annotate uncertainty” option. It is shown on the left side of figure 3 below.

Figure 3. Screenshots of the window for uncertainty parameters of the text.



Target: Value
Certainty level: Unknown
Categories: ignorance, credibility, imprecision, incompletene:
Value:
Description:

Target: Value
Certainty level: Medium
Categories: ignorance, credibility (selected), imprecision, incompletene:
Value: many
Description:

22 For case A, the user selects “Value” from the “Target” combo box and optionally can select the “Certainty level” and “Categories” (from the defined taxonomy of the project). As a result, the selected text is annotated as a segment and the corresponding uncertainty annotation is added in the `<profileDesc>` element. Simultaneously, the annotator (the author) is added to the list of annotators, if missing (see example 2).

23 Case B is a simple extension of A. The user provides an asserted value in the “Value” text input field, which in their opinion should replace the uncertain piece of text (see the right side of figure 3 and the result in example 2).

Example 2. Sample uncertainty annotations made by two users referring to questionable text.

```
<profileDesc>
 […]
 <textClass>
  […]
  <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
```

```
     <certainty xml:id="certainty-A" locus="value" cert="unknown" target="#seg-1"
resp="#annotator-1"/>
     <certainty xml:id="certainty-B" locus="value" cert="medium" ana="https://
providedh.ehum.psnc.pl/api/projects/1/taxonomy#credibility" target="#seg-1"
resp="#annotator-2" assertedValue="many"/>
    </classCode>
   </textClass>
   […]
   <particDesc>
    […]
    <listPerson type="PROVIDEDH Annotators">
     <person xml:id="annotator-1">
      <persName>
       <forename>Michał</forename>
       <surname>Kozak</surname>
       <email>mkozak@man.poznan.pl</email>
      </persName>
     </person>
     <person xml:id="annotator-2">
      <persName>
       <forename>Alejandro</forename>
       <surname>Benito</surname>
       <email>abenito@usal.es</email>
      </persName>
     </person>
    </listPerson>
   </particDesc>
  </profileDesc>
  […]
  <body>
   […] with Donnogh Magwire of Rossbegg in the said County gent, and <seg
xml:id="seg-1">a number of</seg> armed Rebells, who with swords drawen […]
  </body>
```

5.2 Use Case C: Annotating the Entity Associated with a Piece of Text

24 Use case C, along with the remaining types of annotations, concerns the entity types chosen during project creation. For these types of annotations the Annotator uses not `<seg>` but `<name>`. It creates an entity of the provided type, adds it to the list of annotated entities, and annotates the selected text as `<name>` with a `@ref` attribute that points to the created entity. This entity list (or rather these lists, because each entity type has its own list) makes it possible to keep track of named entities in a document and spot possible duplications, and allows opportunities for unification and the easy exploration of the corpus. These lists are added to the `<standOff>` element and are visible to users directly above the text in the Annotator (see figure 2).

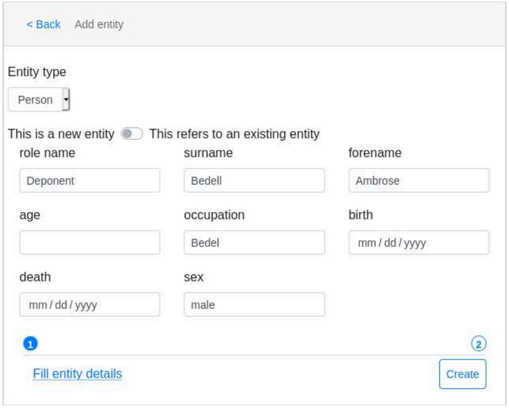
25 This use case focuses on text annotation without uncertainty, illustrated with the following scenario: a user selects “Ambrose Bedell” in the following TEI document and annotates the name as a person.

Example 3. An excerpt of one of the testimonies from the 1641 Deposition data set.

```
  <body>
   […] <pb n="fol. 105r" pagenum="32"/><p>Ambrose Bedell gent sonn to the late
reuerend father in God William Lord Bishop of Kimore in the country of Cavan (by
the crueltie of the Rebells deceased sworne and examined deposeth and sayth […] </
p> […]
  </body>
```

26 The person entity is associated with some predefined properties (predefined subnodes and attributes) that the user can provide. They all are added to the created entity in the dedicated list of annotated entities (in this case people). For instance, from the context the following properties can be extracted: sex, role name, forename, surname, and occupation. Compare the modal window for person properties (figure 4) and the annotation result (`person-1` in example 4).

Figure 4. Screenshot of the window for entity properties.



Entity type: Person

This is a new entity / This refers to an existing entity

role name: Deponent

surname: Bedell

forename: Ambrose

age:

occupation: Bedel

birth: mm / dd / yyyy

death: mm / dd / yyyy

sex: male

Example 4. Sample lists of annotated people from the 1641 Deposition data set.

```
<standOff>
 […]
 <listPerson type="personList">
  <person xml:id="person-1" sex="male">
   <persName>
    <roleName>Deponent</roleName>
    <forename>Ambrose</forename>
    <surname>Bedell</surname>
   </persName>
   <occupation>Bedel</occupation>
  </person>
  <person xml:id="person-2" sex="male">
   <persName>
    <roleName>Victim</roleName>
    <forename>William</forename>
    <surname>Kimore</surname>
   </persName>
   <occupation>Bishop</occupation>
  </person>
  <person xml:id="person-3">
   <persName>
    <roleName>Victim</roleName>
    <forename>Mr</forename>
    <surname>Cloghy</surname>
```

```
     </persName>
    </person>
   </listPerson>
  […]
  </standOff>
```

27 In the text, we use <name> to annotate the selected text and @ref to refer to the entity.

Example 5. Sample annotation of an entity name that refers to the entity.

```
  <body>
   […] <pb n="fol. 105r" pagenum="32"/><p><name xml:id="name-1"
ref="#person-1">Ambrose Bedell</name> gent sonn to the late reuerend father in
God William Lord Bishop of Kimore in the country of Cavan (by the crueltie of the
Rebells deceased sworne and examined deposeth and sayth […] </p> […]
  </body>
```

5.3 Use Cases D and E: Specifying Uncertainty and Alternative Text for Entities Being Annotated

28 These two cases are combinations of the previous ones (A, B, and C) with the same modal windows but shown to the user as a wizard.

29 When annotating with uncertainty a piece of text which becomes a reference to an entity, the corresponding uncertainty annotation points to the added <name>. This allows the user to express their uncertainty regarding the spelling of the entity name. The user does not express any doubts regarding the type or properties of the entity.

Example 6. Sample uncertainty annotations referring to questionable entity name without and with asserted entity name.

```
  <textClass>
   […]
   <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
    <certainty xml:id="certainty-D" locus="value" cert="low" ana="https://
providedh.ehum.psnc.pl/api/projects/1/taxonomy#credibility" target="#name-1"
resp="#annotator-3"/>
```

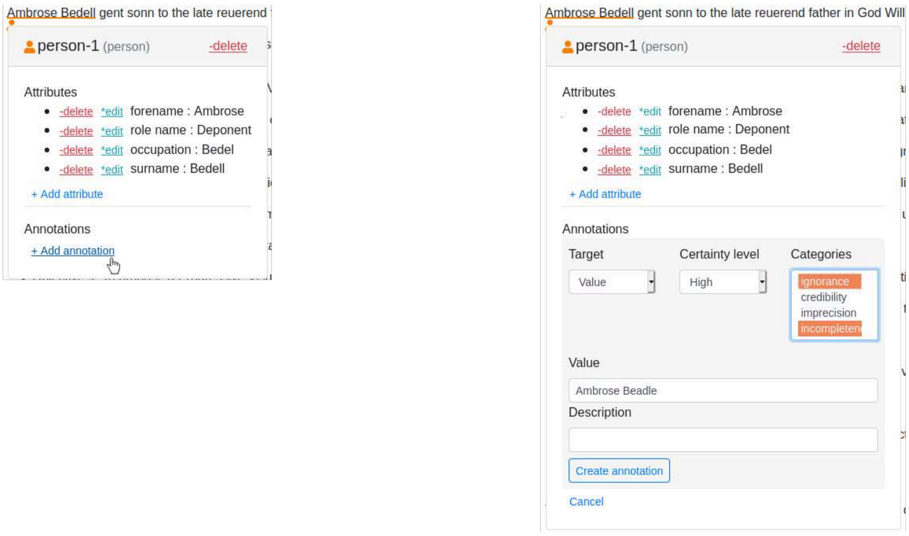
```
    <certainty xml:id="certainty-E" locus="value" cert="medium" ana="https://
providedh.ehum.psnc.pl/api/projects/1/taxonomy#ignorance" target="#name-1"
resp="#annotator-4" assertedValue="#val-E"/>
    <val xml:id="val-E">Ambroze Bedel</val>
   </classCode>
  </textClass>
```

30 It should be noted that, contrary to example 3, the "asserted value" provided by the user consists of more than one word. However, the schema of @assertedValue attribute does not accept whitespaces. In such cases, we propose using the <val> element and pointing to this phrase in @assertedValue. This is also not fully in line with the TEI Guidelines, which only allows identifiers of <anchor> elements (TEI Consortium 2020, ch. 21, Certainty, Precision, and Responsibility).[9] We have, however, submitted a feature request to the TEI-C so that @assertedValue can also take a teidata.pointer value.[10] We use this extension in a number of subsequent cases, to annotate with uncertainty that a piece of text is the name of an entity or that two (or more) entities are in fact the same entity. In these particular cases, @assertedValue must be a pointer, as it takes an attribute value of either @ref or @sameAs (see examples example 10 and example 15). Ultimately, the TEI Consortium is expected to introduce the @assertedValueTarget attribute in a future version of the TEI specification for this purpose. And pointers are to be banned in @assertedValue.

5.4 Use Case F: Annotating an Erroneous Text for an Existing Entity with an Optional Alternative Value

31 This use case is similar to previous use cases D and E, but an instance of the entity already exists in the text (it exists in the dedicated list of entities and its name is annotated in the text). Another user can select the annotated text and add an uncertainty annotation with their doubts and an optional alternative (asserted) text. Modal windows are again presented in a wizard form.

Figure 5. Screenshots of the wizard windows for adding uncertainty annotation.



Ambrose Bedell gent sonn to the late reuerend

person-1 (person) -delete

Attributes

- -delete *edit forename : Ambrose
- -delete *edit role name : Deponent
- -delete *edit occupation : Bedel
- -delete *edit surname : Bedell

+ Add attribute

Annotations

+ Add annotation

Ambrose Bedell gent sonn to the late reuerend father in God Will

person-1 (person) -delete

Attributes

- -delete *edit forename : Ambrose
- -delete *edit role name : Deponent
- -delete *edit occupation : Bedel
- -delete *edit surname : Bedell

+ Add attribute

Annotations

Target: Value

Certainty level: High

Categories: ignorance, credibility, imprecision, incompleten

Value: Ambrose Beadle

Description

Create annotation

Cancel

Example 7. Another sample uncertainty annotation referring to a questionable entity name.

```
  <textClass>
    […]
   <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
    <certainty xml:id="certainty-F" locus="value" cert="high" ana="https://
providedh.ehum.psnc.pl/api/projects/1/taxonomy#ignorance https://
providedh.ehum.psnc.pl/api/projects/1/taxonomy#incompleteness" target="#name-1"
resp="#annotator-2" assertedValue="#val-F"/>
    <val xml:id="val-F">Ambrose Beadle</val>
   </classCode>
  </textClass>
```

5.5 Use Case G: Annotating Uncertainty regarding an Existing Entity’s Type

32 In this use case, the Annotator can annotate a selected fragment even if it crosses hierarchical boundaries in the underlying markup. We encountered such a case in the data set of historical recipes.

Example 8. An excerpt from one of the recipes in the historical recipes data set.

```
  <body>
```

```
   […] Mell vmb, Pache <gap/> sye in butter, das <lb/> sye Zimblich Resch <gap/>
Werden, thue daß <lb/> Gehackte in ein Zinnene <gap/> schisßl <gap/>, Gies <gap/>
<lb/> ein Coppauner […]
  </body>
```

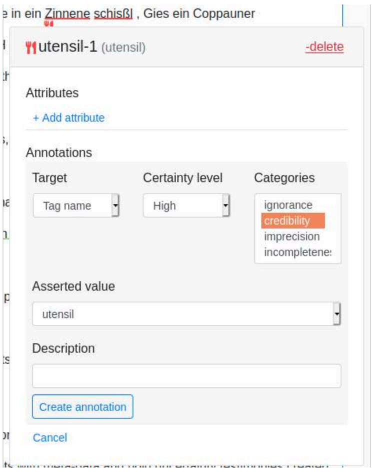
33 Annotating “Zinneneschisßl” as a utensil with a gap inside its name leads us to the following annotations according to use case C:

Example 9. Sample annotation of an entity name divided by a tag.

```
  <body>
   […] Mell vmb, Pache <gap/> sye in butter, das <lb/> sye Zimblich Resch <gap/
> Werden, thue daß <lb/> Gehackte in ein <name xml:id="name-2" next="#name-3"
ref="utensil-1">Zinnene</name> <gap/> <name xml:id="name-3" prev="#name-2"
ref="utensil-1">schisßl</name> <gap/>, Gies <gap/> <lb/> ein Coppauner […]
  </body>
  […]
  <standOff>
   <listObject type="utensilList">
    <object type="utensil" xml:id="utensil-1"/>
   </listObject>
  </standOff>
```

34 If another user would like to express their doubts about the type of this entity, they must proceed as in figure 5 but they must choose “Tag name” in the “Target” combo box. Then the “Asserted value” input text changes to the combo box with defined entity types:

Figure 6. Screenshot of the wizard window for adding uncertainty annotation regarding the entity type.



e in ein Zinnene schisßl , Gies ein Coppauner

utensil-1 (utensil)

Attributes

Annotations

Target Certainty level Categories

Tag name

High

ignorance
credibility
imprecision
incompletenes

Asserted value

utensil

Description

35 The request sent from this modal window is resolved into the following uncertainty annotation:

Example 10. Sample uncertainty annotation referring to questionable entity type.

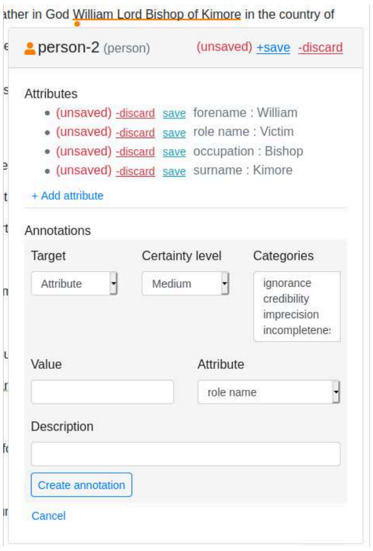
```
  <textClass>
   […]
   <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
    <certainty xml:id="certainty-G" locus="name" cert="high" ana="https://
providedh.ehum.psnc.pl/api/projects/1/taxonomy#credibility" target="#name-2
#name-3" match="@ref" resp="#annotator-1" assertedValue="#utensil-1"/>
   </classCode>
  </textClass>
```

36 As we can see, the “Asserted value” "utensil" is only a request parameter that is resolved in the TEI annotation to the @match attribute with the value "@ref", i.e., the XPath expression that identifies the @ref attributes of two <name> elements. In the TEI annotation, @assertedValue points to "utensil-1" in the list of utensils, but a @locus attribute value of "name" indicates the tag name (i.e., type) of "utensil-1".

5.6 Use Case H: Annotating Uncertainty regarding an Existing Entity’s Property Value

37 Use case H is similar to G. In this case a user wants to directly assign an asserted value to a property of an existing entity about which they have concerns, or simply to express those concerns. In the wizard they have to choose “Attribute” in the “Target” combo box. Then they will be able to provide the property name and an optional asserted value for this property.

Figure 7. Screenshot of the wizard window for adding uncertainty annotation regarding an entity property.



ather in God William Lord Bishop of Kimore in the country of

person-2 (person) (unsaved) +save -discard

Attributes

- (unsaved) -discard save forename : William
- (unsaved) -discard save role name : Victim
- (unsaved) -discard save occupation : Bishop
- (unsaved) -discard save surname : Kimore

Annotations

Target: Attribute

Certainty level: Medium

Categories: ignorance, credibility, imprecision, incompletene:

Value:

Attribute: role name

Description:

Example 11. Sample uncertainty annotation referring to questionable entity property.

```
  <textClass>
   […]
   <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
    <certainty xml:id="certainty-H" locus="name" cert="medium" target="#person-2"
match="persName/roleName" resp="#annotator-2"/>
   </classCode>
  </textClass>
```

5.7 Use Case I: Annotating a Text Fragment as an Occurrence of an Existing Entity

38 Case I is very similar to case C, where we annotate a piece of text which becomes a reference to the entity added to the list. Here we want to annotate that another piece of text refers to the same entity, as in example 12:

Example 12. Sample of a partially annotated fragment of the historical recipe data set.

```
  <body>
   […] Nimbe ein Guetten <name xml:id="name-3" ref="#ingredient-3">Coppaun</name>
<gap/> siede <gap/> <lb/> ihme marb <gap/>, nimb die Prust von dem <lb/> Coppaun,
hack sye garkhlein, Nimb […]
  </body>
  […]
  <standOff>
   <listObject type="utensilList">
    <object type="utensil" xml:id="utensil-1">
     <objectIdentifier>
      <objectName>Zinnschüssel</objectName>
     </objectIdentifier>
    </object>
   </listObject>
   <listObject type="ingredientList">
    <object type="ingredient" xml:id="ingredient-3">
     <objectIdentifier>
      <objectName>Kapaun</objectName>
     </objectIdentifier>
    </object>
    <object type="ingredient" xml:id="ingredient-4">
     <objectIdentifier>
      <objectName>Zimt</objectName>
     </objectIdentifier>
    </object>
    <object type="ingredient" xml:id="ingredient-5">
     <objectIdentifier>
      <objectName>Zitronensaft</objectName>
     </objectIdentifier>
    </object>
   </listObject>
  </standOff>
```

39 This is like use case C except that a user has to toggle the switch from “This is a new entity” to “This refers to an existing entity” (cf. Figure 4) and then select one of the existing entities from the file. The platform resolves the request by adding another <name> tag referring to the same item on the list:

Example 13. Sample name annotations referring to the same entity.

```
  <body>
   […] Nimbe ein Guetten <name xml:id="name-3" ref="#ingredient-3">Coppaun</name> <gap/> siede <gap/> <lb/> ihme marb <gap/>, nimb die Prust von dem <lb/> <name xml:id="name-5" ref="#ingredient-3">Coppaun</name>, hack sye garkhlein, Nimb […]
  </body>
  […]
  <standOff>
   <listObject type="utensilList">
    <object type="utensil" xml:id="utensil-1">
     <objectIdentifier>
      <objectName>Zinnschüssel</objectName>
     </objectIdentifier>
    </object>
   </listObject>
   <listObject type="ingredientList">
    <object type="ingredient" xml:id="ingredient-3">
     <objectIdentifier>
      <objectName>Kapaun</objectName>
     </objectIdentifier>
    </object>
   […]
   </listObject>
  </standOff>
```

5.8 Use Case J: Annotating a Text Fragment as an Occurrence of an Existing Entity with Some Degree of Certainty

40 This case is a simple extension of case I, where assigning the `<name>` to the entity can be annotated with uncertainty.

Example 14. Sample name annotation referring to the same entity with uncertainty.

```
  <textClass>
   […]
   <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
    <certainty xml:id="certainty-J" locus="value" cert="high" target="#name-5"
match="@ref" resp="#annotator-2" assertedValue="#ingredient-3"/>
   </classCode>
  </textClass>
```

5.9 Use case K: Entity Unification

41 Case K describes entity unification, a process that allows specifying with some degree of uncertainty that two entities may be the same. For this use case, we assume that the data set is already annotated manually or automatically, so there are already lists of entities in each TEI file. In a similar manner as for cases I and J, a user can annotate with some degree of uncertainty that two entities are the same. The difference is that we use `@sameAs` instead of `@ref`. For instance when a user unifies a person with the identifier `"person-3"` from example 4 with a person with the identifier `"person-100"` from the file `"file-1"`, then the following uncertainty annotation is created in the TEI file with `"person-3"`:

Example 15. Sample annotation that unifies two entities with some degree of uncertainty.

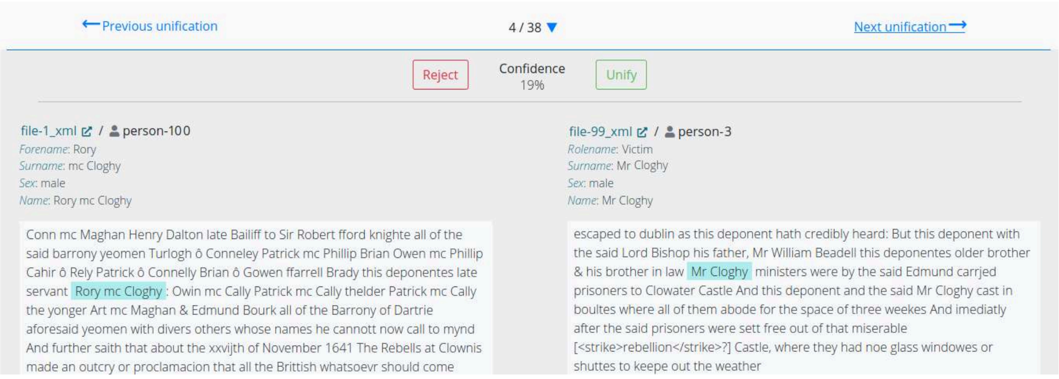
```
  <textClass>
   […]
   <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
    <certainty xml:id="certainty-K" locus="value" cert="high" ana="https://
providedh.ehum.psnc.pl/api/projects/1/taxonomy#imprecision" target="#person-3"
match="@sameAs" resp="#annotator-1" assertedValue="file-1#person-100"/>
   </classCode>
  </textClass>
```

42 Similar but reflexive annotation is created in the file "file-1". In cases when the same user ("annotator-1") had already unified "person-3" or "person-100" with other entities, then the number of annotations will grow exponentially, taking into account that the "sameAs" relation is transitive.

43 Here, it is worth mentioning that these and other annotations are not added directly to TEI files, as we present them above, but they are stored in a database in the form of commits (batches of changes in the project state). These commits are assigned to the file versions they are based on. This way, while rendering the project state to the user, the platform can use XML content from a given file version. In addition, the platform renders the changes from commits associated with this and previous versions, effectively showing to the user the state of the project from any given moment in time.

44 Entities and their properties stored in the database are also the source of data for the automatic unification algorithm we deployed on the platform. It employs a neural network and weighted multigraphs to recognize entities appearing in the data set under different identifiers and to suggest unifications when certain similarities are detected. The platform presents each of the proposed unifications together with the context of the occurrence of the entities being unified, where the user can reject or accept it (see figure 8). In case of acceptance, the user must express their confidence about the unification. The Visual Disambiguator learns from users' actions and subsequent runs of the algorithm take the rejected and accepted unifications into account.

Figure 8. Screenshot of the Visual Disambiguator.



file-1_xml / person-100

Forename: Rory
Surname: mc Cloghy
Sex: male
Name: Rory mc Cloghy

Conn mc Maghan Henry Dalton late Bailiff to Sir Robert fford knighte all of the said barrony yeomen Turlogh ô Conneley Patrick mc Phillip Brian Owen mc Phillip Cahir ô Rely Patrick ô Connelly Brian ô Gowen ffarrell Brady this deponentes late servant Rory mc Cloghy : Owin mc Cally Patrick mc Cally thelder Patrick mc Cally the yonger Art mc Maghan & Edmund Bourk all of the Barrony of Dartrie aforesaid yeomen with divers others whose names he cannott now call to mynd And further saith that about the xxvijth of November 1641 The Rebells at Clownis made an outcry or proclamacion that all the Brittish whatsoevr should come

file-99_xml / person-3

Rolename: Victim
Surname: Mr Cloghy
Sex: male
Name: Mr Cloghy

escaped to dublin as this deponent hath credibly heard: But this deponent with the said Lord Bishop his father, Mr William Beadell this deponentes older brother & his brother in law Mr Cloghy ministers were by the said Edmund carrjed prisoners to Clowater Castle And this deponent and the said Mr Cloghy cast in boultes where all of them abode for the space of three weekes And imediatly after the said prisoners were sett free out of that miserable [<strike>rebellion</strike>?] Castle, where they had noe glass windowes or shuttes to keepe out the weather

5.10 Use Case L: Annotating Uncertainty regarding Other People's Certainty Annotations

45 This last case is all about annotating with uncertainty other uncertainty annotations, a big part of how users can interact with each other in the process of working with TEI files. This enables a process of collaborative disambiguation and editing that starts with a user selecting an uncertain fragment from the list of annotations and choosing "Annotate this." After filling uncertainty parameters ("Certainty level" and "Categories"), the following annotation is created:

Example 16. Sample uncertainty annotation referring to another uncertainty annotation.

```
<textClass>
  […]
  <classCode scheme="http://providedh.eu/uncertainty/ns/1.0">
    <certainty xml:id="certainty-H" locus="name" cert="medium" target="#person-2" match="persName/roleName" resp="#annotator-2" assertedValue="Perpetrator"/>
    <certainty xml:id="certainty-L" locus="name" cert="low" ana="https://providedh.ehum.psnc.pl/api/projects/1/taxonomy#incompleteness" target="#certainty-H" resp="#annotator-1"/>
  </classCode>
</textClass>
```

6. Conclusion

46 Identifying and tracing uncertainty sources and types is an important part of the task of communicating uncertainty in DH research objects, and a process in which the use of the TEI standard can be of great help. Therefore, we identified the need for promoting the use of TEI encoding and providing the tools (both formal and technological) to describe and manage uncertainty throughout the lifecycle of a project. In this paper, we have also presented our approach to creating a taxonomy, the data sets used for its testing, and a collaborative platform that incorporates tools for working with the TEI standard with an emphasis on making uncertainty more present and facilitating collaborative work within the project. We also presented and described how TEI can be used to approach a broad spectrum of use cases where uncertainty can be specified, and how the specification of uncertainty can be modeled in TEI.

47 Finally, another branch of the PROVIDEDH project developed by the Austrian Centre for Digital Humanities and Cultural Heritage is based on annotations of such aspects of historical recipes as ingredients, preparation time, and spiciness. Recipe similarities are calculated based on the co-occurrence of ingredients, their quantities, and their associations. Ingredient unifications are made with the use of mapping tables, where simple and canonical forms of ingredients are preferred rather than fancy mentions like “crisp potatoes,” or “fresh meat.” The project aims to investigate the similarities and interactions between cuisines and, therefore, cultures.[11] This is a good example of using entity annotations and their unifications. Another example of a software system that manages uncertainty (about the location of toponyms in North Africa as they appear in historical sources of medieval and modern times) is described by Martín-Rodilla, Pereira-Fariña, and González-Pérez (2019).

BIBLIOGRAPHY

Binder, Frank, Bastian Entrup, Ines Schiller, and Henning Lobin. 2014. “Uncertain about Uncertainty: Different ways of processing fuzziness in digital humanities data.” In *Proceedings of the Digital Humanities 2014*, 95–98. Accessed December 1, 2020. https://d-nb.info/1164023926/34.

Blau, Adrian. 2011. “Uncertainty and the History of Ideas.” *History and Theory* 50 (3): 358–372. doi:10.1111/j.1468-2303.2011.00590.x.

Drucker, Johanna. 2011. “Humanities Approaches to Graphical Display.” *Digital Humanities Quarterly* 5 (1). Accessed December 1, 2020. http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html.

Edmond, Jennifer. 2019. “Strategies and Recommendations for the Management of Uncertainty in Research Tools and Environments for Digital History.” *Informatics* 6 (3), article 36. doi:10.3390/informatics6030036.

Fisher, Peter. 1999. “Models of Uncertainty in Spatial Data.” *Geographical Information Systems*, vol. 1, *Principles and Technical Issues*, edited by Paul A. Longley, Michael F. Goodchild, David J. Maguire, and David W. Rhind, 191–205. 2nd ed. New York: John Wiley & Sons.

Franke, Max, Ralph Barczok, Steffen Koch, and Dorothea Weltecke. 2019. “Confidence as First-class Attribute in Digital Humanities Data.” In *Proceedings of the 4th Workshop on Visualization for the Digital Humanities (VIS4DH)*. Accessed December 1, 2020. https://vis4dh.dbvis.de/2019/papers/2019/VIS4DH2019_paper_1.pdf.

Hullman, Jessica. 2020. "Why Authors Don't Visualize Uncertainty." *IEEE Transactions on Visualization and Computer Graphics* 26 (1): 130–139. doi:10.1109/TVCG.2019.2934287.

Martín-Rodilla, Patricia, and Cesar González-Pérez. 2019. "Conceptualization and Non-Relational Implementation of Ontological and Epistemic Vagueness of Information in Digital Humanities." *Informatics* 6 (2), article 20. doi:10.3390/informatics6020020.

Martín-Rodilla, Patricia, Martín Pereira-Fariña, and Cesar González-Pérez. 2019. "Qualifying and Quantifying Uncertainty in Digital Humanities: A Fuzzy-Logic Approach." In *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'19).* edited by Miguel Ángel Conde González, Francisco Jesús Rodríguez Sedano, Camino Fernández Llamas, and Francisco José García-Peñalvo, 788–94. New York: ACM. doi:10.1145/3362789.3362833.

Rocha Souza, Renato, Amelie Dorn, Barbara Piringer, and Eveline Wandl-Vogt. 2019. "Towards A Taxonomy of Uncertainties: Analysing Sources of Spatio-Temporal Uncertainty on the Example of Non-Standard German Corpora." *Informatics* 6 (3), article 34. doi:10.3390/informatics6030034.

Simon, Christophe, Philippe Weber, and Mohamed Sallak. 2018 *Data Uncertainty and Important Measures.* London: ISTE; Hoboken, NJ: John Wiley & Sons.

TEI Consortium. 2020. "Certainty, Precision, and Responsibility," sec. 21 in *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 4.1.0. Last updated August 19, 2020.https://www.tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/CE.html.

Therón, Roberto, Alejandro Benito Santos, Rodrigo Santamaría Vicente, and Antonio Losada Gómez. 2019. "Towards an Uncertainty-Aware Visualization in the Digital Humanities." *Informatics* 6 (3), article 31. doi:10.3390/informatics6030031.

Windhager, Florian, Saminu Salisu, and Eva Mayr. 2019. "Exhibiting Uncertainty: Visualizing Data Quality Indicators for Cultural Collections." *Informatics* 6 (3), article 29. doi:10.3390/informatics6030029.

NOTES

[1] Accessed April 25, 2022, http://providedh.eu.

[2] Accessed April 25, 2022, https://providedh.ehum.psnc.pl.

[3] Accessed April 25, 2022, https://www.tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/CE.html.

[4] Online Depositions Website, Trinity College Library Dublin, accessed December 1, 2020, http://1641.tcd.ie/.

5 Rezeptdatenbank der Gastrosophie, History Department, University of Salzburg, accessed December 1, 2020, http://gastrosophie.sbg.ac.at/kbforschung/r-datenbank/.

6 Accessed December 1, 2020, http://gastrosophie.sbg.ac.at/salzburg-zu-tisch/.

7 Zentrum für Gastrosophie, accessed December 1, 2020, https://www.gastrosophie.at.

8 See the ACDH Salzburg Recipes GitHub repository, accessed December 1, 2020, http://github.com/providedh/ACDH_Salzburg_recipes.

9 Accessed April 25, 2022, https://www.tei-c.org/Vault/P5/4.1.0/doc/tei-p5-doc/en/html/CE.html.

10 Michal Kozak, "`@assertedValue` of `<certainty>` should also accept pointers #2067," TEI Guidelines GitHub repository, issue opened December 1, 2020, https://github.com/TEIC/TEI/issues/2067.

11 Renato Souza, food network project in the ACDH-DH "exploration space," GitHub repository, accessed December 1, 2020, https://github.com/exploration-space/food-network

AUTHORS

MICHAŁ KOZAK

Michał Kozak is a designer and programmer of computer systems in the Poznań Supercomputing and Networking Center (PSNC) of the Polish Academy of Sciences. He is a computer science graduate and a doctor of philosophy in Computer Science. Michał received a PhD in computational logic in 2011 from the Faculty of Mathematics and Computer Science of A. Mickiewicz University in Poznań. Since 2010 he has been working for PSNC, initially as a designer and programmer of Java and Python in the Digital Libraries Team, and currently as a leader of the team that creates and develops specialized systems and tools for e-humanities.

ALEJANDRO RODRÍGUEZ

Alejandro Rodríguez Díaz is a research assistant in the Department of Computer Science and Automation at the University of Salamanca. He is a computer engineering graduate from the University of Salamanca, where he is completing his master's degree in intelligent systems. He is currently part of the technical and support staff at the Visual Analytics Group VisUSAL (within the Recognized Research Group GRIAL), where he aids in the design of new visualization techniques for understanding complex data in digital humanities. His areas of interest are data visualization, process automation, and graphic design. Alejandro has special interest in the design of creative and interactive tools for multivariate data.

ALEJANDRO BENITO-SANTOS

Alejandro Benito-Santos is a research assistant in the Department of Computer Science and Automation at the University of Salamanca, which he joined in 2016. Alejandro completed his BSc in Computer Engineering at the same university, from which he also obtained an MSc in Intelligent Systems in 2016. He is a member of the Visual Analytics Group VisUSAL (within the Recognized Research Group GRIAL), where he is currently completing his PhD on text visual analytics (TVA) under the supervision of Dr. Roberto Therón. In his work, he applies visual analytics in a broad range of interdisciplinary research contexts.

ROBERTO THERÓN

Roberto Therón is an associate professor in the Department of Computer Science and Automation at the University of Salamanca. He received the diploma degree in computer science from the University of Salamanca, the B.S. degree from the University of A Coruña, the B.S. degree in communication studies and the B.A. degree in humanities from the University of Salamanca, and the PhD degree from the Research Group Robotics, University of Salamanca. Roberto is currently the Manager of the VisUSAL Group (within the Recognized Research Group GRIAL), University of Salamanca, which combines approaches from computer science, statistics, graphic design, and information visualization to obtain an adequate understanding of complex data sets.

MICHELLE DORAN

Michelle Doran is a postdoctoral research fellow at Trinity Long Room Hub and project officer for the Trinity Centre of Digital Humanities at Trinity College Dublin. She holds a PhD in Medieval Irish Studies, and her principal research interests lie in the field of humanities research and the underlying epistemological and ideological premises. Michelle is the module coordinator of the Digital Scholarship and Skills workshop series hosted by the Trinity Long Room Hub and facilitates a number of workshops on the subjects of digital humanities, data management planning and digital scholarly editing.

AMELIE DORN

Amelie Dorn is a senior researcher at the Austrian Centre for Digital Humanities and Cultural Heritage (ACDH-CH) of the Austrian Academy of Sciences. She received her BA in modern languages and an MA in European Studies from University College Dublin. At Trinity College Dublin she also earned an MPHil and a PhD in Linguistics. Amelie also works as a humanities facilitator in the ACDH-CH “exploration space,” a physical and digital space for innovation and open innovation scenarios. She contributes to various digital humanities projects, and has co-organised a variety of interactions with different actor groups. She also leads the ChIA project, which tests AI and semantic tools for improving cultural image access and analysis on food images from the Europeana database.

JENNIFER EDMOND

Jennifer Edmond is associate professor and co-director of the Center for Digital Humanities at Trinity College Dublin. She holds a PhD in Germanic languages and literatures from Yale University, and applies her training as a scholar of language, narrative, and culture to the study and promotion of advanced methods in, and infrastructures for, the arts and humanities. Jennifer serves as President of the Board of Directors of the pan-European research infrastructure DARIAH-EU. Additionally, she represents this body on the Open Science Policy Platform (OSPP), which supports the European Commission in developing and promoting open science policies.

CEZARY MAZUREK

Cezary Mazurek is a director of Poznań Supercomputing and Networking Center. He received his PhD in Computer Science from Poznań University of Technology in 2004. His expertise and experience is focused on broadly understood ICT applications using research infrastructures. For over twenty-five years of professional activity Cezary has led interdisciplinary teams in which computer scientists and researchers from different domains (e.g., biomedicine, humanities, and earth science) have worked together to address scientific challenges using advanced e-infrastructure services. Recently, he has worked on methods and models for big data processing, the Next Generation Internet initiative (https://www.ngi.eu), and digital humanities.

EVELINE WANDL-VOGT

Eveline Wandl-Vogt is is foundress and orchestra of "exploration space" (2017-) at the Austrian Academy of Sciences. She is foundress and director of the Ars Electronica Research Institute "knowledge for humanity (k4h+)" (2019-) and affiliated to metaLab (at) Harvard. Eveline is an experimentalist, knowledge designer, and digital strategist, working against a background of Art Driven Innovation. She has a multidisciplinary university background, including arts and expertise in knowledge and innovation management. She serves as an expert in various global initiatives, mainly in the area of technical and social infrastructures, and participatory methodologies, such as ADHO, ALLEA, COST actions, DARIAH, and ECSA, as well as standardization bodies. She is an experienced knowledge transfer officer, bridging the gaps between academic knowledge and (social) applications aligned with the SDGs.