Electronic Conservation, Editing and Dissemination of Written Cultural Inheritance

outline for future co-operations between text editors, readers and computer scientists to develop structures and tools to secure and disseminate written cultural inheritance

by Michael Nedo

Conservation of and access to our written cultural inheritance is becoming ever more important. Our trust in progress, which is orientated on goals to be achieved – often solutions to problems resulting from progress – leads not seldom to new, bigger problems: the dilemma of Goethe's 'The Sorcerer's Apprentice'. A reorientation is necessary which has to show us where we are going to and the possible consequences of our progressing. But the orientation of progress can only be determined by the place where we are coming from, by our cultural inheritance, not by the goals science and/or industry are aiming at.

Yet, less and less funding is available to secure, to conserve, to edit and to disseminate cultural inheritance. One reason is that projects of this kind, in particular editorial projects, often become unaccountable, consuming more and more resources over ever-longer periods of time. The 'Leibnitz Edition' for instance, which combines edited text with commentary, was set up before the Great War, it is now expected to be finished in about 200 years with more and more scholars working on it, mostly on its commentary.

Another reason is that politicians demand that printed books, traditional libraries and archives are to be replaced by electronic ones without having contemplated or developed the necessary structures and tools. So far the shelf life of traditional books and documents is rather longer than that of electronic text files which, after about 20 years, become less and less readable, a problem discussed in the US as the 'Big Rot': Are we losing our history?

One further problem is that texts, which represent more than mere scientific or scholarly information, are still more accessible in a printed book than through today's computer technology.

The urgency to secure our cultural inheritance and to make it more accessible requires editorial projects to become accountable and compatible with one another. This requires new editorial methods and suitable electronic technologies:

1. harmonising editorial methods which today are largely idiosyncratic
2. separating conservation and editing of texts from scholarly study by their users: content-related commentary is to be separated from the task of editing the text, it should be created by its users through an interactive communication platform
3. optimising both electronic and printed editions to realise what each does best, so that they supplement each other

These goals can be achieved through collaboration between editors and computer scientists and the scholars and readers working with the editions. The resulting structures and computer tools will finally make editorial projects accountable they will make it possible to produce editions more economically and more faithful, in a better quality. The results of such cooperation will be:

1. computer scientist and mathematicians have to develop structures in which text can be re recorded and archived independent of computer hard- and software, the essential precondition for longevity of text files
2. the development of computer tools to edit, search and sort a text, and the development of an interactive communication platform
3. the development of methods and tools to present and to disseminate text data

An interesting example to demonstrate the advantages and the disadvantages of a printed edition, supplemented by an electronic apparatus, versus a purely electronic edition is the WIENER AUSGABE of Wittgenstein's writings:
The structure of this edition – the microstructure of formal notations and the macro-structure of the page format and of the layout of the WIENER AUSGABE – do not lend them self to be presented on a computer screen, nor is the edition suitable for print on demand. The criteria of readability and of access to the content of a text are still met infinitely better by print, in a well-designed and produced book, than within any computer environment available today. The recognition of this fact finally brought about the change from the originally planned electronic edition of Wittgenstein's work into the printed edition of the WIENER AUSGABE.
Experiments, which were made to improve the quality of proofreading – which was very poor in the computer environment – showed that the quality of proofreading was directly linked to the quality of type in which a text was proofread. This meant that the access to a text – its readability, which allows the concentration needed to study the text – was determined by the quality of type, by the typographical form and by a book-design in harmony with the text. Today this is still only provided by well designed, well printed and produced books, which in turn are only possible based on an underlying, primary electronic editions and suitable computer tools.

What publishers mean today by an electronic edition is usually a searchable, amended and networked PDF/XML version of a printed text, which can also be used for printing on demand. Such an edition is of little use to scholarly research; and printing the WIENER AUSGABE 'on demand' seems to be neither suitable nor economical: in the past the WIENER AUSGABE sold in substantial numbers of copies per volume, largely by subscription, as it will most likely continue to do in the future. The same is true for the student edition of the WIENER AUSGABE. The electronic apparatus supplementing the WIENER AUSGABE will be part of the printed edition, licensed to the subscribers of the edition.

Outline for an electronic apparatus supplementing the WIENER AUSGABE

Rationale of an electronic apparatus is to supplement, not to duplicate a printed edition: Unmediated access to a text, to the author's thoughts is still best provided by a printed edition, which has to be supplemented by powerful electronic search and sort tools.

The Portal to the electronic apparatus of the WIENER AUSGABE is a word (or nominal) concordance accessing the text corpus through a lemmatised dictionary followed by a content (or real) concordance interrelating text content. An example for a word concordance is the *Konkordanz zu den Bänden 1–5* of the WIENER AUSGABE; the *Register zu den Bänden 1–5* is an example for a content concordance. In electronic form these tools can be much extended and improved: printed concordances are only usable if based on a limited text corpus and the search of a word concordance is limited to one word only while words of importance but of high frequency have to be omitted. In the limited size provided in a printed concordance content relations can only be presented as lists.

An electronic apparatus has none of these limitations, it allows the search of word groups and of words in different arrangements: as word combinations like *Raum* and *Zeit* or as occurrences of a word or a group of words within a remark, in groups of remarks, in manuscripts or groups of manuscripts or over the complete text corpus. Research results are initially presented by the frequency of their occurrence, which allows alterations of the search criteria, of search words and their order until the result matches the expectation.
In the electronic apparatus the text is not presented as a facsimile of the printed edition but in a layout and in a typography optimised for the computer screen. The nomenclature of the WIENER AUSGABE is providing a one-to-one relation between the printed text of the edition and its electronic apparatus: volume and page numbers of the edition (which include the references to the underlying MSS and TSS) are followed by the numbers of the remarks, counted for each printed page beginning with 1 and the paragraph numbers within each remark. In the apparatus Wittgenstein's work will be presented either as:

1. the prime text including all open variants and insertions as well as Wittgenstein's markings of and within his philosophical remarks;
2. including Wittgenstein's decided variants, his cancellations;
3. including a variant editor representing the complex structures of Wittgenstein's text variants as well as a representation of the transitions of remarks between manuscripts;
4. including editor's information about text aspects as well as the text genesis;
5. including links to the underlying manuscripts;
6. including links to biographical data and images;
7. including user-commentary linked to their author's

Research results are presented in the context of a remark which can be extended to any number of previous and/or following remarks, these can be further linked to the facsimiles of the underlying manuscripts; additional links can be provided for instance to the posthumous publications by Wittgenstein's heirs, to text translations as well as to content-related commentary and literature.

The word concordance will be supplemented by a concordance of Wittgenstein's formal notations in mathematics and logic and by a concordance of his graphical notations, of diagrams and objects as well as his illustrations.

The prime tool to study Wittgenstein's philosophy is the content concordance. This tool will allow the study of the multitude of interconnections between Wittgenstein's philosophical remarks, between remarks within and between manuscripts; this tool allows the scholar to read Wittgenstein's manuscripts not only in their linear form but also across manuscripts and throughout the whole text corpus, that is around topics, arguments and pictures. Only through this tool does the organic structure of Wittgenstein's work become apparent: individual manuscripts will no longer appear as fragments but as parts of a larger whole, of an organism where every part is related to any other part and where repetitions show themselves as essential to his thinking and writing as he explained in a lecture in 1933:

> There is a truth in Schopenhauer's view that philosophy is an organism, and that a book on philosophy, with a beginning and end, is a sort of contradiction. One difficulty with philosophy is that we lack a synoptic view. We encounter the kind of difficulty we should have with the geography of a country for which we had no map, or else a map of isolated bits. The country we are talking about is language, and the geography its grammar. We can walk about the country quite well, but when forced to make a map, we go wrong. A map will show different roads through the same country, any one of which we can take, though not two, just as in philosophy we must take up problems one by one though in fact each problem leads to a multitude of others. We must wait until we come round to the starting point before we can proceed to another section, that is, before we can either treat the problem we first attacked or proceed to another. In philosophy matters are not simple enough for us to say "Let's get a rough idea", for we do not know the country except by knowing the connections between the roads. So I suggest repetition as a means of surveying the connections.

Part of the electronic apparatus supplementing the WIENER AUSGABE will be an interactive research platform. This will give rise to a lively and creative community of Wittgenstein scholarship and it will create a kind of critical mass of Wittgenstein studies, independent of the locality of the individual reader, scholar or institution. This form of an electronic publication will lead to fast and continuous improvements of the edition, of its apparatus and its commentary while reducing the need for revisions of the printed edition.

A matrix representing internal and external text links of the WIENER AUSGABE

Objects representing Wittgenstein's work are recorded in vertical columns at the top level of the matrix: Wittgenstein's manuscripts (MS) and typescripts (TS) as well as dictations and notes taken by pupils and friends (Nt) and posthumous publications of his work (Pb). Further objects are facsimiles of the originals and biographical ones, i.e. correspondence, travels, images etc., etc.
Horizontal rows of the columns record the internal structures of the objects: beginning with the name of the object followed by its page number (or the number of a cutting in the collections of cuttings) followed by the remark number counted for each page or cutting beginning with one. They are followed by the numbers of paragraphs of a remark and by Wittgenstein's margin signs. For posthumous editions the structures of their editors are included: parts, chapters and sections; Wittgenstein's remarks are counted per page, starting with 1, following the section numbers.
On separate levels of the matrix other objects will be recorded following the basic structure in which Wittgenstein's work is represented, for instance secondary literature and translations based on the posthumous editions. Via an interactive user platform, content related commentary and links researched by users of the WIENER AUSGABE and its Apparatus will be recorded under the names of their authors, as well as mistakes and proposed amendments for the edition and its apparatus.
The network of links will be established on sub-layers of the matrix akin to a multilayer printed circuit. Three types of links will be used: definite, quasi hard wired relations; relations provisionally established, requiring further research and definite verification; and open links still to be established.
The first three layers of links will represent the connections between Wittgenstein's manuscripts and typescripts, at first, a representation of the gestation of the manuscripts over time; second, links between the manuscripts via Wittgenstein's margin signs; and third, the occurrence of identical remarks, i.e. of remarks with identical wording or with identical content albeit changed wording.
External links are presented on further layers: links to posthumous publications and their translations, links to secondary literature, mostly based on published texts, links to content-related commentary under the names of their authors as well as to mistakes in the WIENER AUSGABE and proposed amendments to the edition and its electronic apparatus.
This matrix will be linked via search and sort tools and via an interactive user platform to the text files of Wittgenstein's work. It forms the underlying structure or engine for the content-related or real concordance to Wittgenstein's work.

The development of this matrix and of the tools linking it to the text files of Wittgenstein's work is a joint project of the Wittgenstein Edition at the University of Cambridge and the Department of Computer Linguistics at the University of Munich.