

How to teach your edition how to swim

C. M. Sperberg-McQueen

World Wide Web Consortium/MIT Computer Science and
Artificial Intelligence Laboratory, USA

Abstract

The mutability of electronic editions confronts editors with a new world, in which large parts of current editorial theory must be re-thought, based as it often is on assumptions based on the properties of paper editions. Software can adapt more easily than paper to the needs and interests of the reader, which means many choices about the selection of information in an edition and its presentation to the reader no longer need to be fixed for all time, but can be left open for the reader. Software also tends to have a very short lifetime compared to paper; in order to remain usable for more than a few years, electronic editions must find ways of representing the essential information of the edition in software-independent, non-proprietary ways.

Correspondence:

C. M. Sperberg-McQueen,
259 State Road 399,
Española, NM 87532-3170,
USA.

E-mail:

cmsmcq@acm.org

This essay was written in 1997. It has not been significantly updated but reflects that particular moment in the debate on electronic text.

In recent decades, the electronic preparation of editions has revolutionized the practice of scholarly editing. Now, electronic publication promises a second revolution for editors who have scarcely caught their breath from the first one. This paper outlines what seem to the author to be the most crucial differences between editing for print publication and editing for electronic publication and sketches some demands of the new medium and the best way to meet those demands. Let us begin by considering two pages from important books.

The page in Fig. 1 is carefully designed to give the reader a very specific aesthetic experience. The type was carefully chosen to fit the work; it is laid out on the page so as to suggest the more splendid of the medieval manuscripts in which the work it presents has been preserved. Expansive margins contrast with a dense text block, whose regularity is relieved by intricately decorated initials (Fig. 2). The paper is heavy and well made, although by its nature

the image cannot show this. It is a page made, and exquisitely well made, for looking at.

Now consider, by way of contrast, the pages in Fig. 3.

There is nothing remarkable about this opening; the layout is thoroughly conventional, virtually identical to that found in any of hundreds of editions of the 19th and other centuries. It does not mimic the manuscripts that preserved the poems presented. The typeface used (Fig. 4), though not particularly ugly, is also not particularly beautiful. It was not chosen especially for this edition; many other books published by the same press use the same face. These pages were not made to capture the aesthetic gaze, but as the ample margins and the careful provision of line numbering and poem referencing system show, to be read, re-read, and annotated.

Several differences between these examples are worth pointing out. First, as noted, they differ from each other in numerous details of presentation: font design, type size, page layout, density of the text block, physical size of the page, quality of paper. Second, largely as a result of these differences

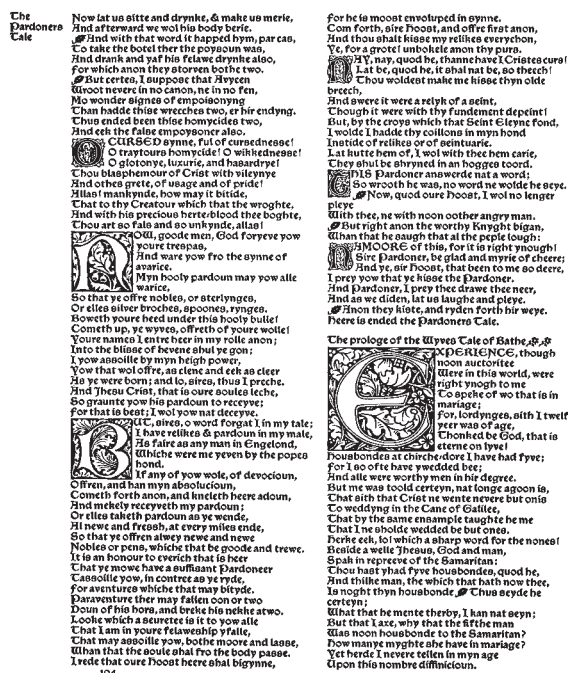


Fig. 1 A page of the Kelmscott Chaucer

of presentation, the two pages have very different effects. The Chaucer edition in Fig. 1 offers the reader an intense aesthetic experience. The archaic text is presented in a luxurious and archaizing typeface and page layout. Although the page appeals in part to the viewer's fantasies of life in Chaucer's England, the experience is offered in some sense ahistorical. No hint of textual problems invades the page; no apparatus of variants reminds the beholder of textual variation and the vagaries of manuscript traditions; no reference numbers suggest the need for commentary or annotation.

The edition of Walther von der Vogelweide in Fig. 3, by contrast, offers itself not primarily as a means to an aesthetic experience but as a functional tool for students and teachers. The apparatus criticus serves as a constant reminder that the text is a historical artifact, the variants as an invitation to consider alternatives to the reading text and to check the editor's work in establishing that text. Any aesthetic encounter with the literary work is the responsibility of the reader; the book design offers little help in the endeavor.

Third, while each is historically important, they are important for very different reasons. The Kelmscott Chaucer, from which the page in Fig. 1 comes, was printed by William Morris at the Kelmscott Press in 1896.¹ It is a milestone in the history of bookmaking in modern Europe. It does not, however, mark any particular achievement in the history of Chaucerian textual criticism. Morris is said, to be sure, to have taken particular care with the text, but by this is meant apparently only that he made inquiries about the available editions and manuscripts and then copied the text out without alteration and without apparatus from Skeat's edition.

The book from which the page in Fig. 3 comes is, by contrast, no monument in the history of Western bookmaking. While not particularly bad, it is also not particularly good. It is mediocre, in both the neutral and pejorative senses. But it is Karl Lachmann's edition of Walther von der Vogelweide.² It would be hard to find another editor whose work on both Middle High German and Latin authors had similar importance in the history of Western text criticism. It has taken Germanic philology more than a century to re-edit all of the Middle High German works edited by Lachmann; more than one editor has spent a career working on a single text which Lachmann edited over the course of a summer or two, only to die with the apparatus of variants still incomplete and still unpublished. Lachmann's Walther (like some others of his Middle High German editions) is still in print for use in lectures and seminars.³

In the context of a collection of essays devoted to the question of electronic editions, however, the most striking characteristics of these two very different editions may be their similarities.

Most obviously, neither edition can conveniently change: each is static, *immutable*. Each chooses one particular view, out of many different possible views, of the text it presents, and one particular way to present that view to readers. Neither edition can change its style of presentation, nor can the information provided be changed conveniently.⁴ Each editor must choose what information about the text to present and how to present it, and each edition is the record of the editor's choices.

As has frequently been observed, electronic publication puts the edition, and thus the editor, into a radically different position. The medium is dynamic and mutable, rather than static or stable. In the

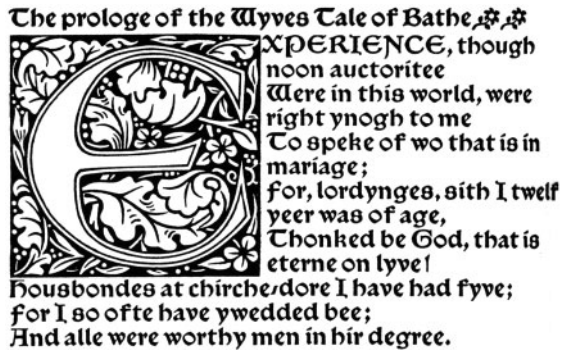


Fig. 2 Closeup of the beginning of the Wife of Bath's Prologue

terminology of computer programming and system design, the print medium forces an *early binding* of the answers to all questions about what information is to be presented to the reader and how it is to be presented, while electronic delivery of editions, by contrast, allows for much later binding. While in print the choices of font, type size, margins, and layout are made and fixed long before the reader takes the book to hand, with electronic delivery it is possible, given suitable planning and software, to delay the choice of font, text size, shape of the text block (window), etc., until the very

Ich saz uf eime steine,
5 und dahte bein mit beine:
dar uf satzt ich den ellenbogen:

Fig. 4 Closeup of the beginning of the *Reichston*

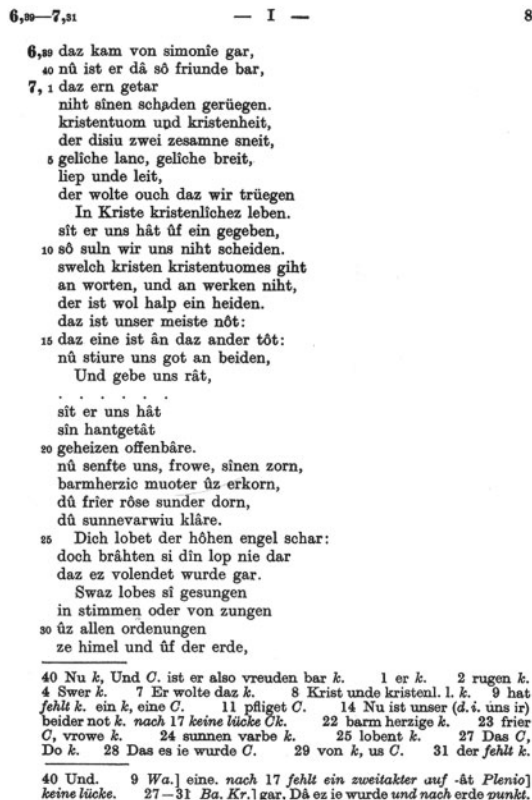
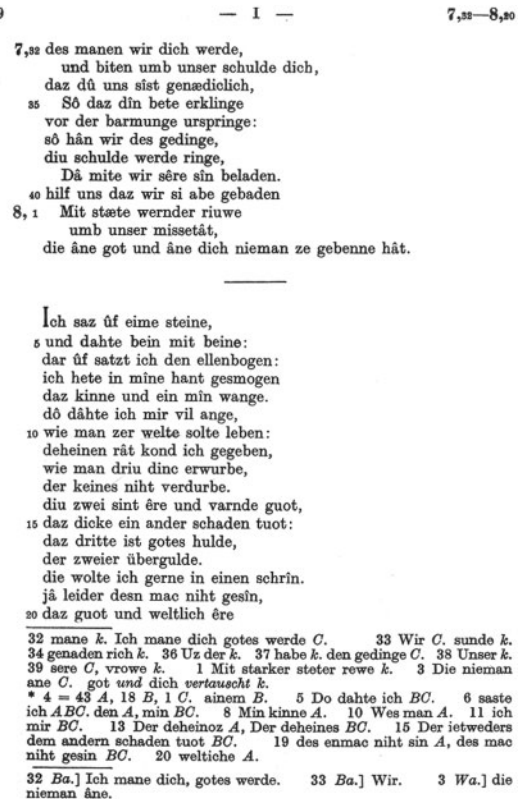


Fig. 3 Two pages from Karl Lachmann's edition of Walther von der Vogelweide



moment of reading, and allow the edition's reader, rather than its producer, to decide on them. By allowing either early or late binding on questions of presentation, electronic editions force their creators to decide when certain choices must be made. If those choices are to be left to the reader, it will be necessary to consider how they are to be made and what alternatives are to be offered.

The possibilities for mutability in electronic editions will require a great deal of editorial theory to be rebuilt. A large part of editorial theory to date has occupied itself with identifying and weighing the relative merits of various choices among alternative positions; these positions have been seen as mutually exclusive, and as requiring an early-bound decision by the editor, rather than a late-bound decision by the reader. But they are mutually exclusive only for static editions; a dynamic edition can allow the reader, rather than the editor, to choose among them. Should the base text be the editor's critical text, or that of a particular witness? Should an apparatus be presented? Which variant readings should be included?

Editors may justly feel that electronic editions have translated them from a stable environment with difficult but well-known problems into a river of Heraclitean flux, in which everything is changing from moment to moment, and the editor and edition are expected to adapt actively to those changes from moment to moment, without being able to rely on many of the principles which used to be stable guides to editorial thinking. Successful engagement with the mutability of electronic editions and the flux of the environment—surviving the experience of being tossed unceremoniously into the river—may require a different way of thinking about editions and the choices they embody.

If editors are to navigate successfully the transition from static to dynamic edition, if their editions are to come to terms successfully with the river of flux into which electronic delivery tosses them, then the creators of editions must learn to understand the nature of the flux and how to use dynamic changes in the edition to deal with it. It no longer suffices to endow an edition with knowledge; electronic delivery demands the *capability* of dealing with mutability. It is in this sense that I mean the reference in

my title. By *teaching an edition how to swim*, I mean endowing an edition not only with a store of factual *knowledge* concerning the work presented, but also with the *capability* of dealing gracefully with the mutability of the electronic medium, by exploiting the possibilities for reader-controlled changes to the edition's presentation and by adapting successfully to rapid changes in the hardware and software environment.

The nature of this knowledge and adaptability can be seen more clearly if we distinguish between two kinds of 'knowing'. The English verb *know* is anomalous among the Germanic languages; its direct cognates have disappeared from the other Germanic languages, but *know* was well established in English by 1200. As the OED says,

The verb has since had a vigorous life, having also occupied with its meaning the original territory of the vb. *wit*, Ger. *wissen*, and that of *can*, so far as this meant to 'know'. Hence Eng. *know* covers the ground of Ger. *wissen*, *kennen*, *erkennen*, and (in part) *können*, of Fr. *connaître* and *savoir*, of L. *novisse*, *co-gnoscere*, and *scire*...

We can distinguish, with the aid of the other languages, several types of knowledge, of which two are relevant here: on the one hand, a knowledge of facts normally transmissible by language (German *Wissen*), and on the other the knowledge of how to do something (German *Können*). Through the former, one may know a great deal about how buoyancy, arm strength, and different types of kick may contribute to efficient swimming: the kind of expert knowledge that Mark Spitz's coaches no doubt possessed. The latter is what Mark Spitz himself knew about swimming: how to actually do it, how to translate thought and knowledge into action. Both kinds of knowledge are essential to the electronic edition—and unlike the print edition, the electronic edition can (and indeed must) incorporate both: the electronic edition can not only know all of the information that is captured as part of its fundamental data, but can also possess information about how that data interacts, how it can be presented, how it behaves.

Knowledge, as I shall use the term here, is a relatively static, immutable body of propositional information. It can be more or less readily paraphrased in declarative sentences. What an edition *knows* is a large body of facts about the work edited, its genesis and/or transmission, etc.

Capability, by contrast, is the capacity to *do* things. It is intimately connected to dynamic phenomena, to the way things happen in time, and thus to mutability. What an electronic edition *can do* will involve interaction with the reader, adjustment to user options, and the like. Textually, it may be expressed in a series of imperative sentences. In practice, capability is intimately connected with the execution of software, and as both theoretical computer science and several decades of practical software development soberly attest, capability is not necessarily subject to paraphrase.

Given a suitable *declarative* notation for knowledge, that is, it is in principle possible and usually straightforward as a matter of practice to translate the knowledge into another notation. Given an executable program, it is in principle not universally possible to recognize what it does without running the program, and in practice it is notoriously difficult to understand executable code or to replicate it in a different notation. The difference is important because in practice some information can be captured either in a declarative or in an imperative formulation. The declarative form lends itself to paraphrase and thus tends to be more software- and hardware-independent, while the imperative form tends to be tightly bound to specific pieces of executable software and to resist successful paraphrase.

It will be useful to distinguish three things, which in the following discussion I will refer to simply as the *facts*, the *selection*, and the *presentation*.

- There is an infinite set of *facts* related to the work being edited. Some are directly observable, some are not directly observable but are recoverable by inference or intuition, and some are irrecoverable.
- Any edition records a *selection* from the observable and the recoverable portions of this infinite set of facts. This may be thought of as the paraphrasable information content of the edition,

or the set of inferences we may make from the edition. An edition may include an apparatus of variants or not, may provide glosses for some words or not, may provide historical or literary annotation or not, etc. Each of these different styles of edition represents a different selection from the brute facts.⁵

- Each edition provides some specific *presentation* of its selection: the text is given in a particular layout and font, the apparatus if any is given at the bottom of the page (or window) or at the back of the book (in another window), variant readings are signaled by Nestlé symbols or not,⁶ glosses are presented in the margin or in a free-standing glossary, etc.

The *selection* and *presentation* of an edition together constitute what may be called its *interface*: the face it presents to its reader.⁷

Mutability (or dynamic change) is the fundamental condition of the electronic delivery of editions; it is visible at several levels. The *facts* from which the edition selects are themselves in some sense mutable; this is true for all editions, not only for electronic editions. New manuscripts of a work will sometimes come to light. Different scholars, especially scholars from different generations, will assess the relevance and the recoverability of various facts differently. Bédier revolutionized the editing of Old French, for example, by his declaration that reliable manuscript stemmata are not reconstructible, and that the text of the archetype, and a fortiori that of the author, are thus unrecoverable.⁸ Later editors have in some cases moved the argument further, taking the view that even if the archetype were reconstructible, it is less important for the study of the work than a synoptic view of the entire textual tradition in which the work was disseminated. This observable fluctuation in scholarly views of the relevant and available factual record is surely responsible in part for the gradual decline in the idea of a definitive edition. The immutability of the print edition makes it hard for print editions to represent or adapt to the mutability of the facts.

At the level of the interface, the *mutability* inherent in electronic delivery makes it possible to provide cinematic visual effects, atmospheric sound, music, film clips of readings or performances, and

the like. As more and more such opportunities are seized, the edition more and more resembles a movie or television presentation. The exploitation of synchronized multimedia can make the work more accessible to readers and help them come more quickly to grips with the aesthetic nature of otherwise difficult and inaccessible literary works.⁹ The edition, or the software in which it is presented, can further allow for certain kinds of user interaction which affect the presentation: users may be allowed to change the font or font size (this is particularly useful in making editions more accessible to readers with visual handicaps),¹⁰ or even the layout and color scheme of the screen presentation.

Good interactive design, it must be observed, is a new challenge to which interface designers have not yet fully adapted. The first instinct of many designers, when confronted with the extreme mutability of windows-based user interfaces, is to seek ways to limit the reader's freedom in adjusting the interface; when they succeed, the reader confronts windows which cannot be resized, fonts which do not react to requests to enlarge or reduce them, and a presentation that tends to treat the screen as a slightly inferior form of paper and the reader as a passive viewer rather than an active explorer. User interfaces to electronic editions are more rewarding when their designers recognize that the design challenge is not to assume a fixed window size and layout and create a pleasing static screen based on those assumptions, but to create a design which adjusts well and automatically as the reader changes the window size and aspect ratio, enlarges the font, etc. It is not formally a requirement that editions provide interactivity or multimedia design which adapts to the reader's changes in this way, but editions which do not will inevitably seem wooden and stiff to readers accustomed to well-designed interactive multimedia presentations; editions will be more attractive to readers, and more usable, if they can swim.

Even more significantly than its presentation, an edition's *selection* from the facts has historically been determined largely by the editor's choices about the audience and purpose of the edition. The Kelmscott Chaucer and Lachmann's edition of

Walther discussed above illustrate two different kinds of selectivity, as well as two different kinds of presentation. Electronic delivery opens new possibilities here: the selection of information need not be fixed once for all by the editor, but can be mutable and changed dynamically by the reader. If the edition includes an apparatus of variants, for example, the software's user interface can allow the user to decide whether to display the apparatus or not, and how to display it. Variant readings can be displayed in a traditional apparatus (possibly in any of several conventional styles, positive, negative, with different punctuation) in a separate window or a separate frame, or in a synoptic display of the witnesses in parallel columns or parallel horizontal lines, or in a variety of other ways.¹¹ If the edition's selection from the set of brute facts provides the necessary information, it is possible to allow the reader

- to read any witness in its entirety, from beginning to end
- to read (in a genetic edition) the editor's reconstruction of a particular version in the genesis of the text, in clear text, whether that genetic version is identical to the text offered by any documentary witness or not
- to examine the genesis of the work, as hypothesized by the editor, by means of a dynamic illustration of textual layers¹²
- to jump from the transcription of any witness to its photographic reproduction
- to shift from any point in one witness or version of the work to the corresponding location in any other version
- to have two or more versions on the screen scrolling in parallel
- to display any witness, version, or editorial reconstruction we choose as the base text, and have an apparatus in traditional form at the bottom of the screen or in a separate window, which scrolls along in sequence with the base text
- to have two or more witnesses on screen, with a common apparatus keyed to either one, or with an apparatus for each
- to make the apparatus selective by witness (showing only selected manuscripts) or by type of variation (e.g. suppressing all purely orthographic variants)

- to show the entire tradition in line-by-line synoptic form
- to use special symbols (e.g. symbols similar to those used in the Nestlé Greek New Testament) to provide a visible signal in the base text of entries in the apparatus, or to provide a base text clear of such intrusions
- to show the text in old-spelling or new-spelling form
- to display glosses in the margin, or at the bottom of the window, or in a separate window, with or without markers on glossed words to signal the availability of a gloss
- to display annotation (or markers indicating the availability of notes) selectively by type: glosses, explanations of linguistic difficulties, text-critical notes, notes on historical background, literary analysis, etc.

Note that the capabilities offered by the presentation layer of an electronic edition depend in part upon the selection of facts represented in the edition. And conversely, an editor's selection of information to record in the edition may be influenced by the interface foreseen for the edition. The user can control whether and how glosses are presented only if the edition provides glosses, and similarly for other forms of annotation, variant readings, normalized spelling, and the reconstruction of versions representing stages of the work's genesis.

Because delivery software can filter the body of information contained in the edition at the time of reading, the same edition can be used to generate the forms of presentation thought most useful to beginning students, to advanced students, or to specialists with various interests. It is thus no longer necessary to choose once for all between the aesthetically oriented, apparatus-free presentation embodied by the Kelmscott Chaucer and the functional presentation illustrated by Lachmann's Walther. Nor is it necessary to perpetuate the not wholly satisfactory compromises characteristic of some recent American scholarly editing (include the apparatus, but present it in the least convenient way possible, separated from the base text but in the same volume, so that text and apparatus simply *cannot* be viewed together without buying two copies). It is possible to provide both views of

the work, with and without the apparatus, in the same electronic edition. A multi-purpose edition can be created simply by including all the information needed for each purpose (typically there will be more or less overlap) and providing a suitable user interface for each.

Many editors will find the dynamic filtering of information a siren song, re-awakening in them the dream of an ultimately complete edition, which contains everything a reader of the work might need or wish to consult: medievalists, classicists, and Bible scholars may find themselves dreaming about editions with full transcriptions of every textual witness, with full paleographic detail; modernists may set about the construction of archival editions with photographic reproductions of every authorial draft and every edition with any possible claim to authority; literary historians may indulge themselves in providing full-text access not only to the work at the center of their edition, but to its literary sources, potential sources, analogues, and to literary works for which it, in turn, has served as a source.¹³ It may be worth stressing, therefore, that however we enlarge the selection of facts recorded in an edition or archive, it remains inescapably a selection. Electronic delivery mechanisms do not impose quite the same tax on additional material as does book printing, and dynamic filtering of the content can help keep the reader from being overwhelmed by the volume of material. But it should be noted that making an edition capable of serving multiple purposes may require more additional effort than first appears. A well-done study edition of an important work is not simply a scholarly edition minus the apparatus: it will require a different apparatus of its own, with annotations and glosses (and perhaps a new-spelling text) suitable for the audience. Even in the electronic world, it appears, the editor will need to make choices, and the edition will be a record of those choices. Absolute completeness remains a chimera.

The underlying structural fact about editions which are capable of serving multiple purposes is that they successfully separate the data—their knowledge of facts—from any specific behavior based on that data, and they separate both from the specific machinery and tools which present this

information to the reader. This separation turns out to be crucial not only for this reason, but also because it allows the edition to stay afloat in yet another kind of flux. It is not only the user interface and the selection of facts presented in the edition which are mutable: the hardware and software substrate used in electronic delivery of editions are themselves mutable, and their mutability threatens the longevity of electronic editions.

Scholars have learned recently to beware of acid paper and the potential that books printed on it will fall apart with use and age unless active steps are taken to preserve them. To forestall this problem, printed scholarly editions are nowadays printed on paper designed to last 300 years or longer without chemical deterioration. But if they are tied to a specific piece of software, electronic editions risk the equivalent of publication on paper so acidic that it decomposes in 15 or 20 years. Both the hardware and the software marketplaces change quickly: today's hit machine is tomorrow's slowpoke; today's standard software is tomorrow's rarity. Some early literary hypertexts are now in the gradual process of becoming unreadable, as the hardware/software configuration for which they were originally written become rarer. Hypercard, which was released for the first time in 1986, was last updated in 1998; although as I write this Hypercard readers are still available, they are incompatible with some newer hardware, so that Hypercard stacks are unrunnable on some current machines. IRIS Intermedia, the first version of which became available in 1985, was orphaned in 1992 when the Brown University Institute for Research in Information and Scholarship, which developed it, was shut down. The DynaText browser, which combined strong hypertext capabilities with the ability to work with SGML-encoded documents, was effectively on the market for only a decade or so. For editorial projects which hope to produce editions which will serve readers for more than a few years, and which themselves may have lifetimes measured in decades, it is clearly dangerous to tie the edition too closely to a single piece of software or hardware.

To survive for more than a few years, an edition must thus be able to survive in the constant flux of hardware and software. There is of course no

absolute requirement that an edition be able to swim in this sense. Editors seldom believe nowadays in a 'definitive' edition, nor hope to create an edition which will never need to be replaced. An edition, it is often now said, is the reflex of a particular point of view, at a particular time. New times, we say, will always require new editions. It is thus possible, at least in theory, to view with equanimity the prospect of an edition becoming inaccessible to readers because its software layer no longer runs on currently available computer systems. It does not matter, we might say, that the edition is no longer available. It *did* once exist, at the time for which it was intended. And after its time has passed, surely (we might ask) there is no point in trying to extend its lifetime by artificial means?

In practice, it would seem that most editors may well wish for the first fascicle of their electronic edition still to be working when they issue the final fascicle; some may hope that their editions will still be usable on the day the editor retires. Editions made for the moment may also be of interest after that moment has passed, not perhaps as the most reliable guide to the work, but as a guide to the reception of the work, or to the moment for which the edition was created.¹⁴ Editors who harbor such hopes will need to be careful in the organization of their editorial work, because experience shows clearly that the average longevity of text browsing software does not meet such demands. Note that it is not a matter of making a better choice of software, or of devising a superior user interface for it: what is required is that the edition strictly limit its dependency on any single piece of software, and that the creators of editions plan for the obsolescence of their delivery systems.

If, as I have suggested above, it is the mutability of electronic editions which poses the greatest challenge for the would-be creators of electronic editions, then are there specific steps that editors can take in order to meet that challenge in all its various forms? There are. The remarks which follow are not exhaustive, but will provide, I hope, useful guidance for the perplexed.

Editors can and should exploit the mutability or dynamism of the edition's interface for editorial purposes: an edition which responds interactively

to the reader and provides information in multiple media can provide a more persuasive aesthetic experience of the work, as well as a more useful tool for study. The literature on user interfaces may usefully be consulted.¹⁵ Existing user interfaces, both for editions and other text browsing/display applications, and those of other software applications, may provide inspiration for rethinking the interface of the edition. In addition to user interfaces in the strict sense, the larger field of information design may repay consideration.¹⁶

Electronic editions can serve multiple audiences much more conveniently than print editions can. Success in this endeavor will require that the editor carefully analyze the kinds of information and the styles of presentation needed by different readers or for different purposes. Existing principles of editorial theory should in principle be helpful here, but some parts of current theory are tainted by silent (and perhaps unconscious) assumptions based on the immutability of print editions.¹⁷

The information content of the edition should be recorded in declarative terms, as *knowledge*, independent of the expected presentation, avoiding reference to presentation details and imperative phrasing. The selection of information to be captured in the edition will be affected not only by the editor's evaluation of the documentary record but also by the intended interface. If, in the course of preparing an edition or a particular interface, one decides unexpectedly that certain lemmata or variant readings should be displayed in red, one may need to revise the encoding of the text to ensure that those lemmata or variants are distinguishable from the others in some way. This is a straightforward piece of document analysis: the fact that the decision comes late in the process does not change its nature. It would be wise, because more generally useful, to ask oneself what it is which sets these readings apart from the others, and to tag them accordingly, instead of adding a *red* tag to the encoded text. In a different interface, those readings may be colored blue instead of red; a third may be designed for color-blind readers who need a different form of visual signal.

The *form* in which the edition records information will also be affected by the intended

presentation. Users of XML markup not infrequently find themselves having to choose among several possible ways of expressing the same information; this is a function of the richness of practical markup vocabularies and the existence of multiple points of view from which a given piece of information may be understood. When facing such a choice, one might choose a particular way of encoding the information in order to ensure that the delivery software can produce a particular desired presentational effect with it. If the several methods of expression are truly synonymous with each other,¹⁸ then there is, in principle, no objection to making the choice on these grounds. It is a serious mistake, however, to allow one's expected delivery software to distort one's understanding of the text, or one's encoding scheme.¹⁹

The textual information, however, can and should be distinguished rigorously from its intended presentation. Being able to provide a declarative description of this information, without reference to presentation, is a useful test of whether the information has been kept sufficiently distinct from presentational issues.²⁰

It is desirable (although the practice is neither widespread in electronic editions nor predominant in software development) to provide more than one interface to the edition. In software development terms, the interface is taken care of by a *front end*, while the information presented resides in a *back end*. Providing multiple interfaces to the edition amounts, in these terms, to providing more than one front end to the same back end. Multiple front ends have two potential advantages. Most obviously, they can support different ways of working with the text, different cognitive styles in the reader or student, different interests. One interface might focus on the aesthetic experience of the work, another on its literary-historical background, another might focus on issues of textual variation or textual genesis, another on linguistic questions. It may be possible in theory to provide access to all of these different subject matters from the same interface, but (as illustrated by Morris's Chaucer) some forms of interface benefit if they are not obligated to refer the reader to other possible interfaces but can concentrate on providing a self-contained experience.

Multiple interfaces can also help the editor ensure that the core information in the edition is in fact recorded declaratively and without excessively tight binding to any particular style of presentation. When different interfaces are designed to present the same information in different ways, it is easier to focus on the common declarative non-presentational interpretation of the information and easier to detect failures to separate information from presentation.²¹

Within any single interface, general principles of user interface design should be applied. Default behaviors should be set to cater to the novice or the first-time user of the edition. Some aspects of the user interface (e.g. some of the more arcane user-settable options) will need to remain unobtrusive so that they do not confuse the novice. They should not, however, be made completely inaccessible unless this is intrinsic to the design of the specific interface: it should be possible for readers who have become very familiar with the edition's interface and have specialized requirements to exploit the mutability of the edition's interface to the fullest, so that they can meet those requirements.

It is probably desirable, although in the current state of computing knowledge difficult, to separate the design of the presentation layer of the editions interface into a declarative description of the presentation desired and a separate software layer which provides the actual interface to the reader, guided by the declarative description. Note that this amounts to capturing as much of the presentation layer as possible in declarable, expressible *knowledge*, leaving as little reliance on pure *capability* as possible. In practice, this division is today only partially possible: stylesheets can and should be used to ensure device independence of the delivery mechanisms and to capture as much presentation information as possible in declarative form. Current stylesheet formalisms are notoriously weak, however, in providing declarative descriptions of interactive behaviors: they are much better developed for static aspects of presentation than for the dynamic aspects.

The separation of information from its presentation, and the provision of multiple interfaces, will be

substantially easier if the core information of the edition in encoded is a declarative, non-procedural form. SGML (the Standard Generalized Markup Language) or its subset XML (the Extensible Markup Language) are obvious possible choices, although some of their crucial advantages can also be achieved with other markup schemes. The Perseus Project and Peter Robinson's *Canterbury Tales* project both encoded their textual material in SGML and translated it into non-SGML form for delivery with specific software at the time of issue. As the software packages initially used inevitably fall out of favor and into disuse, the editions themselves can readily be re-issued in newer software. The Wittgenstein Archive at the University of Bergen similarly practices a sharp distinction between the core of the edition and its presentation. The project devised its own markup scheme, the Multi-Element Code System (MECS), and markup vocabulary (MECS-WIT) for the transcription of Ludwig Wittgenstein's notebooks. Because the markup is declarative and independent of any particular style of presentation, it supports a variety of interfaces, from diplomatic transcriptions of the notebooks with precise attention to details of the inscription to normalized-spelling fair-copy displays of the text.²² Unless there is a good reason to do otherwise, it is probably advisable to use an existing publicly documented markup vocabulary for editions, such as that developed by the Text Encoding Initiative (TEI), or the TEI extension developed by the Model Editions Partnership (MEP), or a further local adaptation of these.²³

Non-textual data should be archived in non-proprietary, publicly documented data formats; these have much greater longevity than privately controlled formats. If necessary, the data can be translated into a proprietary format for delivery with specific software, thus introducing a sharp distinction between the *archival* and *delivery* forms of the edition; this is in any case a useful distinction to make. Use of non-proprietary formats has as a side effect that it will normally be easier for the editor or others to integrate new material into an edition, thus ameliorating some of the drawbacks associated with the need for a fixed and immutable selection of core content in an edition. An *open* edition can be

hospitable to the incorporation of new material and thus adapt more successfully to the mutability of the documentary and factual record.

Scholarly editions have historically had the task of documenting and correcting for the mutability of texts as transmitted from author through copyist or typesetter to reader. There is perhaps an irony in the observation that in order to serve their purposes in a new medium, they must not resist but must embrace and learn to exploit mutability for their own purposes. Some editors will surely find it distressing to discover that they must endow their editions not only with a solid basis of factual knowledge, but must also give them the capability of *doing* things interactively with the reader.

The times are changing. We had better start swimming, or we'll sink like a stone.

Notes

The talk from which this essay was derived was given under the title 'Why you should not teach your edition how to swim'; after some soul-searching, and with some misgivings, I have retained the metaphor, which seems to have stuck in the auditors' minds, but reversed its polarity. The argument has benefited considerably from discussion at the Conference on Editorial Problems, and in particular from the comments of Julia Flanders, François LaChance, and Peter Robinson. The author is grateful for encouragement and advice received, during the revision for this volume, from Julia Flanders and from Fred Unwalla, whose patience has been beyond all saying.

- 1 The images are reproduced from *The Works of Geoffrey Chaucer: A Facsimile of the William Morris Kelmscott Chaucer* with the original 87 illustrations by Edward Burne-Jones, together with an introduction by John T. Winterich. Cleveland and New York: World Publishing Company, 1958, p. 104.
- 2 The images are reproduced from *Die Gedichte Walthers von der Vogelweide: Studienausgabe* herausgegeben von Karl Lachmann; dreizehnte, aufgrund der zehnten, von Carl von Kraus bearbeiteten Ausgabe, neu herausgegeben von Hugo Kuhn. Berlin: de Gruyter, 1965, pp. 8–9.
- 3 An informative summary of Lachmann's work as an editor of Middle and Old High German works can be found in Hendricus Sparnaay, *Karl Lachmann als Germanist*. Bern: Francke, 1948.
- 4 Two exceptions to this characterization may be worth noting. At the cost of some effort any reader can add marginal or external annotations and variant readings. And in the case of the Walther edition, or any edition with an apparatus of variants, a sufficiently energetic reader can reconstruct an alternative text, e.g. an approximation of the text offered by each manuscript witness. Such an exercise involves a great deal of to and fro between the base text and the apparatus, which gives the base text a psychological prominence which may be out of keeping with its textual importance. Since it remains easier and more convenient to read the base text than to reconstruct an alternative text from the apparatus, any edition with a single base text and apparatus can hardly escape biasing its readers toward its base text, however that base text is arrived at by the editor and whatever the claims the editor may make or avoid making for it. Here, the selectivity of Lachmann's apparatus means that for the witnesses to Walther only an approximation to the text of any given witness is possible in any case.
- 5 Some readers may wonder what to make of the term *fact*, which seems to ascribe somewhat more certainty to the contents of editions and the judgments of their editors than is always warranted. An editor may be undecided which variant reading to prefer, or uncertain how to read a smudged page, or hesitant about the meaning of a hard place; surely we should not ignore these cases when characterizing the contents of the edition? I am taking the term *fact* here with a rather large denotation and do not mean, by using it, to exclude cases of uncertainty from our consideration. If one wishes to insist that the term *fact* necessarily involves a claim of certainty, then we may perhaps reconcile ourselves to the term by observing that in the cases described it is a wholly certain fact that the editor is uncertain in the ways described, the nature of the arguments seen to bear on the case is also a fact, and editions can in fact record both the uncertainty and the arguments on either side of such a decision. In any case, nothing in the argument relies on the 'facts' recorded in an edition having any particular degree of certainty; on the contrary, the certainty attributed to the facts or propositional content of an edition is itself mutable, and it is somewhat easier for electronic editions to come to terms with this mutability than for print editions.
- 6 These are graphic symbols appearing in the base text, which call attention to the presence of variants in the apparatus; different shapes are used to signal one- and multi-word variants, additional text in other versions, text missing in other versions, and transpositions. Their name derives from their use in Eberhard and

- Erwin Nestlé's editions of the Greek New Testament, although similar symbols have also been used in New Testament editions at least since the 1880s.
- 7 I am indebted to Julia Flanders for the suggestion that it could be useful to generalize from the software notion of user interfaces to a broader concept which includes the analogous features of print editions.
 - 8 **Bédier, J.** (1928). *La tradition manuscrite du Lai de l'Ombre*: Reflexions sur l'art d'éditer les anciens texts. *Romania*, 54: 161–196, 321–356.
 - 9 Michael Groden's edition *James Joyce's Ulysses in Hypermedia* (project home page at <http://publish.uwo.ca/~mgroden/ulysses/>) is a well known project which works hard to exploit the opportunities offered by synchronized multimedia presentation. It is suggestive that Groden is identified not as an editor but as the 'Director' of the project.
 - 10 A full account of accessibility issues is not possible here; the interested reader should consult, among others, the World Wide Web Consortium's guidelines on this topic: *Web Content Accessibility Guidelines 1.0*, **Chisholm, W., Vanderheiden, G. and Jacobs, I.** (eds), W3C Recommendation 5 May 1999 (Cambridge, Sophia-Antipolis, and Tokyo: World Wide Web Consortium, 1999), available online at <http://www.w3.org/TR/WAI-WEBCONTENT/>; *Web Content Accessibility Guidelines 2.0*, **Ben Caldwell et al.** (eds), W3C Working Draft 22 August 2002. Cambridge, Sophia-Antipolis, and Tokyo: World Wide Web Consortium, 2002, online at <http://www.w3.org/TR/WCAG20/>; *User Agent Accessibility Guidelines 1.0*, **Jacobs, I. Gunderson, J. and Hansen, E.** (eds), W3C Proposed Recommendation 16 October 2002. Cambridge, Sophia-Antipolis, and Tokyo: World Wide Web Consortium, 2002, online at <http://www.w3.org/TR/UAAG10/>; *XML Accessibility Guidelines*, **Dardailler, D., Palmer, S. B. and McCathieNeville, C.** (eds), W3C Working Draft 3 October 2002. Cambridge, Sophia-Antipolis, and Tokyo: World Wide Web Consortium, 2002, online at <http://www.w3.org/TR/xag>. General information may be found at the home page of the W3C Web Accessibility Initiative (<http://www.w3.org/WAI/>).
 - 11 For proposals on the display of variants, see (inter alia) **Steer, G.** (1968). Grundsätzliche Überlegungen und Vorschläge zur Rationalisierung des Lesartenapparats. In Kuhn, H., Stackmann, K. and Wuttke, D. (eds), *Kolloquium über Probleme altgermanistischer Editionen. Marbach am Neckar, 26. u. 27. Apr. 1966. Referate und Diskussionsbeiträge*. Wiesbaden: Franz Steiner Verlag, pp. 34–41, and for an argument against the use of apparatus see **Ijsewijn, J.** (1977) *Companion to Neo-Latin Studies*. Amsterdam, New York, Oxford: North-Holland, chapter 4.
 - 12 Jean-Louis Lebrave's prototype genetic edition of Flaubert's *Hérodias* provides a vivid illustration of possibilities in this vein; see **LeBrave, J.-L.** (1991). 'L'hypertexte et l'avant-texte' in *Texte et ordinateur: les mutations du lire-écrire : Actes du colloque interdisciplinaire tenu à l'Université Paris X Nanterre, 6–8 juin 1990*, sous la direction de Jacques Anis et Jean-Louis Lebrave. La Garenne-Colombes: Éditions de l'Espace Européen, pp. 101–117; **LeBrave, J.-L.** (1997). 'Hypertexte et édition génétique: l'exemple d'Hérodias de Flaubert,' in *Banques de données et hypertextes pour l'étude du roman*, sous la direction de Nathalie Ferrand. Paris: Presses Universitaires de France, pp. 137–154; and **Zeller, Z. and Martens, G.** (eds) (1998). 'Hypertext und textgenetische Edition' in *Textgenetische Edition*, Tübingen: Max Niemeyer Verlag, pp. 329–45. Another useful exploration of the problems of electronic genetic editions, with more emphasis on practical issues and their solutions, can be found in **Rischer, T.** (1997). 'Eine TEI/SGML-Edition der textkritischen Ausgabe von James Joyce *Ulysses*'. Diplomarbeit, TU München.
 - 13 The endeavors described here in the conditional are not fanciful. The history of computing in the humanities is full of projects whose salient character is their relative completeness in comparison with their pre-electronic forebears. In addition to Roberto Busa's *Index Thomisticus* (a full concordance, without stop words, of the complete works of Thomas Aquinas, as well as all the works once erroneously attributed to him), and the *Thesaurus Linguae Graecae* (a transcription of the entire extant corpus of ancient Greek), projects more explicitly identified as editorial in nature can readily be named here. Hoyt Duggan's projected edition of *Piers Plowman*, described in **Duggan, H. N.** (1996). 'Some unrevolutionary aspects of computer editing', in Finneran, R. J. (ed.), *The Literary Text in the Digital Age*, Ann Arbor: University of Michigan Press, pp. 77–98 (and see the project's home page at <http://www.iath.virginia.edu/seenet/piers/piers-main.html>), and Peter Robinson's *Canterbury Tales* edition, illustrate the reaction postulated of medievalists. Jerome McGann's Rossetti Archive, described in 'The Rossetti Archive and image-based electronic editing', in *The Literary Text in the Digital Age*, ed. Finneran, pp. 145–183, is a prominent example of the archival impulse. As examples of the literary-historical desire to accumulate massive annotation and collections of parallel texts, one might mention Michael Groden's project for a hypertextual Joyce edition

- (project home page at <http://publish.uwo.ca/~mgroden/ulysses/>) and the print and CD-ROM edition *Der junge Goethe in seiner Zeit: Sämtliche Werke, Briefe, Tagebücher und Schriften bis 1775*, Eibl, K., Jannidis, F. and Willems, M. (eds). Frankfurt a.M., Leipzig: Insel, 1998.
- 14 The patristic editions of J.-P. Migne, for example, are no longer regarded as the best editions for the study of the Church Fathers: the major authors, and many of the others, are now available in more reliable editions. The editions of Migne, however, remain a milestone in the history of patristics, and an important source for the history of the reception of patristic thought. They may no longer be important as editions, but they remain important as historical monuments in their own right.
 - 15 A good place to start is Borenstein, N. S. (1991). *Programming as if People Mattered: Friendly Programs, Software Engineering, and Other Noble Delusions*. Princeton: Princeton University Press, and the literature cited there.
 - 16 By far the best place to start in this topic is the work of Edward Tufte: Tufte, E. (2001). *The Visual Display of Quantitative Information*, 2nd edn Cheshire, CT: Graphics Press, Tufte, E. (1990). *Envisioning Information*. Cheshire, CT: Graphics Press, and Tufte, E. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, CT: Graphics Press. Although he focuses on quantitative information, his example can be inspiring also for those concerned with the display of textual and other non-numeric information.
 - 17 An outstanding, and in some respects alarming, example is offered by Tanselle, G. T. (1972). Some Principles for Editorial Apparatus. *Studies in Bibliography*, 25: 41–88, rpt. in his *Textual Criticism and Scholarly Editing*. Charlottesville: University Press of Virginia, 1990, 19–176. Almost none of the principles Tanselle identifies apply to editions with mutable interfaces.
 - 18 For a discussion of markup interpretation which provides a framework within which notions like synonymy can usefully be applied, see Sperberg-McQueen, C. M. Huitfeldt, C. and Renear, A. (2001). Meaning and Interpretation of Markup *Markup languages: Theory & practice*, 2.3: 215–234. <http://www.w3.org/People/cmsmcq/2000/mim.html>
 - 19 On these and other issues relating to markup design, see the growing literature on SGML, including especially Maler, E. and d'Andaloussi, J. (1996). *Developing SGML DTDs: From Text to Model to Markup*. Upper Saddle River, NJ: Prentice Hall PTR.
 - 20 Some works may require special treatment: the poet Stefan George insisted that his poetry be presented only in sans-serif type, never with serifs.
 - 21 Borenstein, *Programming as if People Mattered*, suggests writing software with multiple user interfaces for similar reasons. 'If software provides multiple user interfaces from the start, it is hard to avoid making the code separation clean and complete' (p. 47). In addition, he notes that providing multiple interfaces makes the gradual evolution of software and of user interfaces over several releases easier for users to accept and adjust to: it gives software users more control over their use of the software. The same arguments, *mutatis mutandis*, apply to editions and their interfaces.
 - 22 See Huitfeldt, C. (2000). *MECS – A multi-element code system*, in *Skriftserie fra Wittgensteinarkivet ved Universitetet i Bergen*. Bergen: WAUB, <http://www.hit.uib.no/claus/mecs/> and Wittgenstein, L. (1998–1999). *Wittgenstein's Nachlass: The Bergen Electronic Edition*. CD-ROM edition in four volumes. Oxford: Oxford University Press, 1998–99.
 - 23 For the TEI vocabulary, see Sperberg-McQueen, C. M. and Burnard, L. (2002). Text Encoding Initiative, *Guidelines for Electronic Text Encoding and Interchange* (TEI P4), XML-compatible edition prepared by Syd Bauman, Lou Burnard, Steven DeRose, and Sebastian Rahtz. Oxford, Providence, Charlottesville, Bergen: TEI Consortium. The MEP extension of the TEI vocabulary is documented in Chesnutt, D. R., Hockey, S. M. and Sperberg-McQueen, C. M. (1999). Markup Guidelines for Documentary Editions, working paper of the Model Editions Partnership, 4 July 1999. Columbia, SC: Model Editions Partnership, <http://adh.sc.edu/MepGuide.html> and Chesnutt, D. et al. (2002). Model Editions Partnership reference manual: Tag set documentation. Columbia, SC: Model Editions Partnership <http://adh.sc.edu/meptsdvl.html>.