

Leeds Studies in English

Article:

Oliver S. Pickering, 'The Computerisation of the Index of Middle English Prose: The Way Forward?', *Leeds Studies in English*, n.s. 27 (1996), 163-71

Permanent URL:

https://ludos.leeds.ac.uk:443/R/-?func=dbin-jump-full&object_id=123743&silos_library=GEN01



Leeds Studies in English
School of English
University of Leeds
<http://www.leeds.ac.uk/lse>

The Computerisation of the *Index of Middle English Prose*: The Way Forward?

O. S. Pickering

The *Index of Middle English Prose (IMEP)*, the international enterprise to catalogue and index all surviving examples of Middle English prose on the basis of their manuscripts, has achieved what can be described as 'critical mass'. Eleven handlists have so far appeared, describing and indexing a total of some 3,900 items in almost 750 manuscripts: a large and important body of material, even though the project is still much nearer its beginning than its end.¹ But although size – and a growing professionalism of presentation – has brought *IMEP* respectability, it has also created a major disadvantage, which will only increase as each new handlist adds to the accumulating data.

The disadvantage is that of fragmentation. An index, to be easily usable, should be single; with *IMEP* every new volume adds to the disunity. The resulting frustration is felt not only by those who try to use the series for their research but by those involved in the preparation of future handlists. The stage has been reached where, save for major or easily identifiable prose works, it has become a time-consuming and not wholly foolproof task to establish whether or not any of the published volumes has already dealt with an item.² This is not to take away from the real achievements of the handlists, many of which represent lasting works of scholarship. But the indexes of *incipits* and *explicitis* which routinely conclude each volume are useful only so far as scribes may be trusted always to have used the same form of words. The accompanying 'general' indexes suffer from inconsistent subject terminology, and their replacement in some recent volumes by separate indexes of authors, titles, persons / places, and subjects – only increasing the number of different places in which to search – has not improved matters.

The general editor has worked hard to impose order on his contributors, but everyone involved in *IMEP* – a project which will in time wholly transform our knowledge of Middle English prose – cannot escape the disadvantage that 1978, the

year of the founding conference, was really too early: the computer had not yet arrived. Those who set up the project were correct to realise that the work of listing, identifying and codifying all extant items of Middle English prose – never before attempted – was so large and complex that it would have to proceed repository by repository, with the production of a cumulative index thus delayed; this decision would surely have been the same if work had started in 1988.³ There is no doubt also that printed handlists were the right way to proceed, and that they remain today the best means of presenting the fruits of individual scholars' long and intimate involvement with the collections assigned to them.

But the creation of a cumulated *IMEP* – combining the findings of the published handlists – can be seen now to be an urgent task that cries out for computerisation, not only because of the fragmentation described above but because the organisers did not sufficiently foresee the severe methodological problems that have arisen as a result of different scholars encountering the same prose material in different manuscript collections. Indexers are asked to make reference to other known manuscripts of items they come across. If a work was frequently copied, has been edited, and is bibliographically 'secure', they are rightly urged to refer the reader to the *Index of Printed Middle English Prose (IPMEP)*⁴ rather than list the same (say) fifteen manuscripts each time. But if the work in question is little-known (and was little-copied) and the indexer has been able to add to knowledge by sorting out the manuscript situation, then the other manuscripts are of course to be listed.

These practices may seem straightforward, although indexers have not always found it easy to maintain a balance between self-sufficiency, in a reference book sense, and unnecessary duplication of information. But there is the complicating factor that time and scholarship march on, represented in part by the publication of the indexers' own labours: the present total of eleven *IMEP* handlists, produced in the course of as many years (1984-95) by different scholars who have not only frequently followed somewhat different procedures (partly because indexing technique has naturally evolved) but who have had access to different states of bibliographical knowledge. For many prose items this means that both *IPMEP* and certain of the handlists are already out of date bibliographically, with the result that further unqualified cross-reference to them may be misleading. A later handlist may instead become, for a time, the new benchmark of information about an item, its author having been able to supplement or correct the previous received wisdom. This is a natural and encouraging development, but the fact of the new authority is difficult to communicate to the indexer-in-progress let alone the innocent researcher, partly because it has not been the norm for indexers to refer routinely to earlier handlists for occurrences of items already

indexed. In consequence the perpetuation of superseded scholarship remains a real possibility. Indexers, publishing their results piecemeal, are in effect undermining existing bibliographical authority without being enabled to contribute to a new order.

In 1978 A. S. G. Edwards anticipated that it would be 'at least two decades before a definitive *Index of Middle English Prose* can appear'.⁵ Three or four decades would now seem more realistic, given the state of academic life, but slow progress means that the computer can be brought to bear at what is still a relatively early stage of the enterprise. For it is surely the computer which will solve (or at least greatly reduce) the methodological crisis which is now in danger of overwhelming *IMEP*.

If and when computerisation is applied to the project – and it may be that this cannot happen without major funding being forthcoming – it will be an opportunity to rethink the present procedures from scratch, not least the way in which the data itself is organised. Cumulating the published handlists into a single database, searchable by powerful free-text (or field-based) methods and made available either online or on CD-ROM, is an obvious step and not what is most at issue here. The advantages of being able to identify items of prose by textual elements other than their alphabetically-arranged opening or closing lines, to search rapidly for items of prose on particular subjects, and to search for information that is not at present indexed at all, such as manuscript locations and editions of Middle English prose, are not in question. Similarly, the possible search techniques, which might include both keyword searches operating on automatically-created indexes and 'direct text' searches for Middle English words or phrases, displaying the results in context, are no more than might now be expected, given the state of technology. However, it may be noted that the existing data would very likely be enhanced, during computerisation, to make the retrieval of information easier: one necessary enhancement would be the provision of consistent subject-indexing terms, while another would be the introduction of selective modernisation of the Middle English text (alongside the original), possibly in the dual form of a short, modernised form of *incipit* and *explicit* together with significant, modernised 'text keywords' selected from the full *incipit* and *explicit* (see the example below).

In the light of the discussion that began this paper, the more significant benefit afforded by computerisation would be the opportunity to revise, standardise and reorganise *IMEP* data so as to create, for the first time, an authoritative master database of bibliographical information about Middle English prose. Indeed the combined cumulation and reorganisation of the published data, while the enterprise is still in progress, would in practice achieve at a much earlier stage than would otherwise be possible the original declared aim of producing a merged *IMEP* able to

stand alongside the *Index of Middle English Verse*.⁶ It would also significantly hasten the completion of the printed handlists.

The very fact of cumulation would necessitate the *revision* of the existing data, at the very least, for a cumulative *IMEP* would be a single work in which the user would naturally expect consistency: the tolerance that may be allowed to individual handlists, produced at different times, would not be so easily granted to a whole. Given the discrepancies in the handlists to which attention has been drawn above, the computerisation of the data should be seen as an opportunity to produce, in effect, a second edition of the existing *IMEP*: a database which will transform the existing raw data into an authoritative source of information about Middle English prose by reorganising the data into what may be called 'master records' and 'copy records'.

Under this radically new division of data, the master records would (it is proposed) hold a small amount of textual information in modernised form, but would mainly provide information relating to the item as a 'composition', separate from its manuscript manifestations. In contrast the copy records would contain information relating to the particular manuscript copies of each item, one such record for every manuscript, and would include the full Middle English *incipit* and *explicit* as in the printed handlists. Hypertext links would enable immediate cross-reference between copy and master records. The user of the database would be able to choose between searching master records (say, by subject term) and copy records (say, by phrases of Middle English text). The distinction between master and copy records would mean that 'global' information about any one item of Middle English prose – including lists of its manuscripts and printings – would be held once only, removing at a stroke the problems of 'authority' that beset the existing handlists.

Thus a master record might contain the following fields of information (I preface each field with a two-letter code, solely for the purposes of the present illustration):

- MI Brief *incipit*, modernised
- ME Brief *explicit*, modernised
- AU Modern form of author's name, if known
- MT Modern title and/or brief description of the work
- KW Selected text keywords, modernised
- SU Subject indexing terms
- BR References to standard works of bibliographical reference
- ED Editions
- CR Hypertext links to copy records for manuscripts already indexed in *IMEP*

- OM Other known manuscripts
MN Unique record number

In contrast, a copy record might well contain the following fields:

- MS Manuscript and folio reference
HA *IMEP* handlist / page / item reference
CI Middle English *incipit* (fifty words)
CE Middle English *explicit* (twenty words)
CT Middle English title or rubric
EM Editions from the manuscript in question
TX Textual peculiarities, and references to published discussion of the manuscript text in question
MR Hypertext link to master record
MD References to descriptions of the manuscript
NO Other manuscript notes
CN Unique record number

This distinction between master and copy records may be demonstrated by the following example taken from the forthcoming *Handlist of Middle English Prose in the Manuscripts of Lambeth Palace Library*, now in an advanced stage of preparation.⁷ The printed version is expected to read much as follows:⁸

MS 432

Described James and Jenkins, pp. 599-601.

[11]

f. 90

The right gloriouse virgyn seint dorothee come down of the noble blode of the senatours of rome hir fadir hight dorotheo and her modir hight theodera in that tyme percecucioun of crysten peple was passing grete in the londe of romaynis wherfore this blessid dorotheo dispising the ydolis forsoke rome . . .

f. 93v

. . . and coeternall with the fader and the holy goest lyvith and reynyth god by all the worldis of worldis iblessid amen

Life of St Dorothy, followed by Latin text and a final prayer, ending (f. 94) 'quod fuller'. *IPMEP* 696, Wells Rev. 2: V [75d], in both as if unique, but Görlach, pp.

34-40, shows that Wells Rev. is wrong to classify versions 75c-e separately: all preserve the same translation from Latin, which was later revised into one of the 'additional legends' inserted into the *Gilte Legende* (see the entry, above, for Lambeth 72 [92]). Ed. from the present manuscript in Horstmann, 'Prosalegenden', pp. 325-28.

Other texts: BL Royal 2 A xviii, f. 236v, and Manchester Chetham's 8009, f. 1, the latter indexed in *IMEP* II, pp. 86-87. The *Gilte Legende* revision is preserved in BL Add 11565, Add 35298, and Lambeth 72.

MS s. xv med., by a single scribe signing himself (Richard) Fuller. For a suggestion that it was made at Syon Abbey, see Keiser, 'Patronage and Piety', p. 43.

Under the suggested computerised model this data would be enhanced and redistributed into a master and a copy record, as follows. For the sake of the example the forthcoming Lambeth handlist is referred to here as *IMEP* XV; the page number within it is invented as are, of course, the unique record numbers.

Master record

- MI The right glorious virgin saint dorothy . . .
ME . . . by all the worlds of worlds blessed amen
AU
MT *Life of St Dorothy*, translated from Latin and later revised into one of the 'additional legends' inserted into the *Gilte Legende*
KW virgin; dorothy; blood; senators; rome; dorotheo; theodora; persecution; christian; romans; idols
SU Saints; Hagiography; St Dorothy
BR *IPMEP* 696, Wells Rev. 2: V [75c-e], the latter wrong to classify these three versions separately, as all preserve the same translation from Latin (see Görlach, pp. 34-40)
ED Ed. from Lambeth Palace 432 in Horstmann, 'Prosalegenden', pp. 325-28.
CR Manchester Chetham's 8009 (II.174), Lambeth Palace 432 (XV.195)
OM BL Royal 2 A xviii, f. 236v
MN 1234

Copy record

- MS Lambeth Palace Library MS 432, ff. 90-93v
HA *IMEP* XV, p. 123 [11]
CI The right gloriousse virgyn seint dorothee come down of the noble blode of the
senatours of rome hir fadir hight dorotheo and her modir hight theodera in that
tyme percecucioun of crysten peple was passing grete in the londe of romaynis
wherfore this blessid dorotheo dispising the ydolis forsoke rome . . .
CE . . . and coeternall with the fader and the holy goest lyvith and reynyth god by
all the worldis of worldis iblessid amen
CT
EM Ed. in Horstmann, 'Prosalegenden', pp. 325-28.
TX Followed by Latin text and a final prayer, ending (f. 94) 'quod fuller'
MR 1234
MD James and Jenkins, pp. 599-601
NO MS s. xv med., by a single scribe signing himself (Richard) Fuller. For a
suggestion that it was made at Syon Abbey, see Keiser, 'Patronage and Piety',
p. 43
CN XV.195

The above master record would be shared with other manuscript copies of this *Life of St Dorothy*, while the copy record would by its nature be unique. If and when there were scholarly developments affecting the item as a whole the master record alone would need updating, whereas under the printed model every related handlist entry would become out of date.

The amount of work needed to transform the printed *IMEP* handlists into a database of this kind should not of course be underestimated. There is little doubt that recasting the existing handlist entries into master and copy records could occupy a well-qualified research assistant for two years, making a considerable amount of funding a pre-requisite. And before such a researcher could start there would be the two tasks of software development – no small matter with data of *IMEP*'s complexity – and of keyboarding or otherwise converting to electronic format those printed handlists which are not already available on disk. Finally there would need to be a means of keeping the database up to date (making corrections in the light of new research, adding material from newly-published handlists) once the retrospective conversion had been completed.

It may be that some colleagues might question the usefulness of creating both

master and copy records in every single case, for example for minor items such as recipes, which are so often unique and which are less susceptible to bibliographical control. But it is the present view of the *IMEP's* Computerisation Sub-Committee (which includes the author of this paper) that a system similar to that described above, rigorously applied, is the only intellectually coherent way of organising the data emerging from the handlists so as to realise *IMEP's* full potential to benefit future scholarship. The wealth of data is such that this benefit would undoubtedly be enormous.⁹

NOTES

¹ The handlists so far published (all by D. S. Brewer, Cambridge) cover the following collections: I, The Huntington Library (Ralph Hanna III, 1984); II, The John Rylands Library and Chetham's Library (G. A. Lester, 1985); III, The Digby Collection in the Bodleian Library (Patrick J. Horner, 1986); IV, The Douce Manuscripts in the Bodleian Library (Laurel Braswell, 1987); V, Additional Manuscripts 10001-14000 in the British Library (Peter Brown and Elton D. Higgs, 1988); VI, Yorkshire Libraries and Archives (O. S. Pickering and Susan Powell, 1989); VII, Parisian Libraries (James Simpson, 1989); VIII, Oxford College Libraries (S. J. Ogilvie-Thomson, 1991); IX, The Ashmole Collection in the Bodleian Library (L. M. Eldredge, 1992); X, Scandinavian Collections (Irma Taavitsainen, 1994); and XI, The Library of Trinity College, Cambridge (Linne R. Mooney, 1995). The general editor of the series is A. S. G. Edwards.

² Cf. Jeremy Griffiths's review of Handlists V-VIII in *The Library*, 6th series, 18 (1996), 56-58: 'it is now quite a performance to search for an item of Middle English prose in the indexes of all these volumes' (p. 58).

³ The deliberations of the founding conference are published in A. S. G. Edwards and Derek Pearsall, ed., *Middle English Prose: Essays on Bibliographical Problems* (New York and London, 1981).

⁴ R. E. Lewis, N. F. Blake and A. S. G. Edwards, ed., *Index of Printed Middle English Prose* (New York and London, 1985).

⁵ A. S. G. Edwards, 'Towards an Index of Middle English Prose', in Edwards and Pearsall, *Middle English Prose*, pp. 23-41 (p. 26).

⁶ A revision of the *Index of Middle English Verse* is also currently being undertaken, by Julia Boffey, A. S. G. Edwards, and Linne R. Mooney.

⁷ By O. S. Pickering and V. M. O'Mara.

⁸ *IMEP* handlists naturally make use of abbreviated forms of frequently-cited works. The references here to James and Jenkins, Görlach, Horstmann, and Keiser are of this type, and need not be explained for the present illustrative purpose. 'Wells Rev.' is the revised version, in progress since 1967, of *A Manual of Writings in Middle English*, first compiled in 1916 by J. E. Wells.

⁹ The Sub-Committee would welcome comments on the proposals outlined in this paper.