

## The Domesday Texts Project

Kumamoto University, Japan, 24 May 2018

### Introduction

Domesday Book is undoubtedly the most famous medieval document in the world. A Google Alert regularly provides twenty to thirty new references to the source each week from around the world. Such global recognition is in large part due to the power of the name rather than any awareness of or interest in its contents. Domesday, with its reference to the day of judgement, resonates in the same way as, say, The Tibetan Book of the Dead does. It is invoked to enhance vicarious authority and authenticity. For historians, however, it is a fundamental source for English history. Compiled from a survey of landed income and resources in 1086 by reference to tenure in 1066, it documents the Norman Conquest of England, a key turning point in the nation's story.

For the academic and interested layperson alike Domesday Book is best known today from the modern translations in the Phillimore and Alecto editions.<sup>1</sup> Completed in 1986 and 2002 respectively, both draw on modern scholarship. The only text we have of the original Latin, however, was produced by Abraham Farley in 1783.<sup>2</sup> It is generally agreed that this text is extremely accurate for its time, indeed even a wonder of scholarship today. There is something of a fan club out there that resists any suggestion that it is inadequate. But inadequate it is. With the benefit of the high resolution facsimile of the Digital Domesday published by Alecto,<sup>3</sup> it is now clear that Farley contains many more errors of transcription than hitherto appreciated. More importantly, it is not a critical edition that meets the demands of modern scholarship. I am currently engaged in re-editing Great Domesday Book (GDB), the first volume of this iconic work, that covers all of England less the eastern counties of Norfolk, Suffolk, and Essex, to meet those demands.

### Expansion of the text

What, then, is new about the edition? Strictly speaking, the Farley text is not an edition at all. A special typeface, known as Record Type, was cast to represent all the forms of the manuscript (figure 1). So, Farley is more accurately a print facsimile as if were; its aim was to reproduce exactly what the scribe wrote. This, though, often obscures what the scribe intended to convey. In accordance with medieval practice, he abbreviated his words to speed up the process of writing. The abbreviations marks in themselves are valuable evidence. John Palmer has recently shown that their forms and distribution reflect the sources that the scribe drew on.<sup>4</sup> But they are difficult to read and often hide nuances of meaning. So, the new edition extends the Latin to remove ambiguities.

---

<sup>1</sup> Phillimore: *Domesday Book*, ed. J. Morris and others, 34 vols, London 1974–86, available online as *The Electronic Edition of Domesday Book: Translation, Databases and Scholarly Commentary, 1086*, 2nd. ed., at <https://hydra.hull.ac.uk/resources/hull:domesdayTranslation> or <http://discover.ukdataservice.ac.uk/catalogue?sn=5694>; Alecto: *Domesday Book*, ed. R.W. H. Erskine, A. Williams, and G. H. Martin, London 1986–2002.

<sup>2</sup> *Domesday Book; seu Liber Censualis Willelmi Primi Regis Angliae inter Archivos Regni in Domo Capitulari Westmonasterii Asservatus*, ed. A. Farley, 2 vols, London 1783.

<sup>3</sup> *The Digital Domesday*, CDRom, Alecto Historical Editions, London 2002.

<sup>4</sup> J. J. N. Palmer, 'A Digital Latin Domesday', in *Domesday Now: New Approaches to the Inquest and the Book*, ed. David Roffe and K. S. B. Keats-Rohan, Woodbridge 2016, 61-80.

**TERRA PICOT DE GRENTEBRË IN STANES HVNDRET**  
**P**icot de Grentebrige tenet In COELA .iiii. hidas  
 7 dimidiam. et x. acras. Terra. est. v. carucis. In dominio. est. una. 7 viii.  
 uillani. habent. iiii. carucas. Ibi. ii. molini 7 dimidius de. xxi. i. solidis.  
 Pratum. carucis. Inter. totum. ualeit. et. ualuit. vii. libras. Hanc. terram  
 De. hac. terra. tenuerunt. Alricus. <sup>monachus</sup> et. Godric. iiii. hidas. et. dimidiam.  
 homines. abbatis. de. Ramesey. non. potuerunt. recedere. absque. licentia. eius.  
 et. iiii. sochemanni. homines. Regis. Edwardi. tenuerunt. i. hidam. et. x. acras. non. potuerunt  
 dare. nec. uendere. sine. licentia. regis. IN. CILDEFORD. HVNDRET.  
**I**n. Pampesuoorde. tenet. Radulfus. de. picot. iiii. uirgas. Terra. est  
 .i. carucis. et. ibi. est. et. ii. acrae. prati. Valet. et. ualuit. semper. x. solidos. Hanc  
 tenuit. Edericus. homo. Alurici. cilt. et. recedere. potuit. cum. terra.  
**I**pse. Picot. .... tenet. HESTITONE. IN. WITELESFORD. HVNDRET.  
 Pro. xv. hidis. et. dimidia. se. defendit. Terra. est. xiii. carucis. In. dominio. vii.  
 hide. et. iiii. uirgae. et. ibi. est. .i. caruca. et. iii. adhuc. possunt. fieri.  
 Ibi. xx. uillani. cum. xii. bordariis. habent. ix. carucas. Ibi. ii. molini  
 de. xxi. solidis. et. iiii. denariis. Pratum. iiii. carucis. In. totis. ualentis. ualeit  
 x. libras. Quando. recepit. xvi. libras. et. tantundem. T. R. E. Hanc. terram  
 recepit. picot. pro. iobus. maneris. ut. dicit. Viginti. sochemanni. tenuerunt.  
 Horum. unus. homo. algari. <sup>comitis</sup> dimidiam. hidam. habuit. non. potuerunt. recedere  
 Alii. uero. homines. Regis. Edwardi. fuerunt. et. recedere. potuerunt. et. viii. aueras  
 et. viii. Inewardos. et. iii. heuewardos. uicecomiti. inueniunt.

Figure 1: a. the Farley transcript of part of GDB 200, the beginning of Picot of Cambridge's chapter, b. expansion of the same entries.

The process is revealing some interesting characteristics. When one dips into Farley there is a tendency to assume that the scribe was always competent. He was not. He often got tired and his Latinity - his fluency in Latin - deserted him. In places, for example, it is common to find a singular subject and plural verb, or vice versa. What was intentional and what was accidental becomes an issue and introduces uncertainty. Conversely, new insights emerge. For example, the formula *silua x porc'* has always been understood as *silua x porcis*, 'woodland for so many pigs', and thus has been seen as a unambiguous measure of woodland. However, a *punctus elevates*, an upside-down semicolon, is regularly inserted after *silua*, indicating that a phrase is to be understood (figure 2). From the small number of cases in which it is supplied, it is clear that the pigs were the render that tenants made for the right to pannage, the pasturing of their pigs. Since rates varied from area to area and probably village to village, the data afford no straightforward measure of woodland.

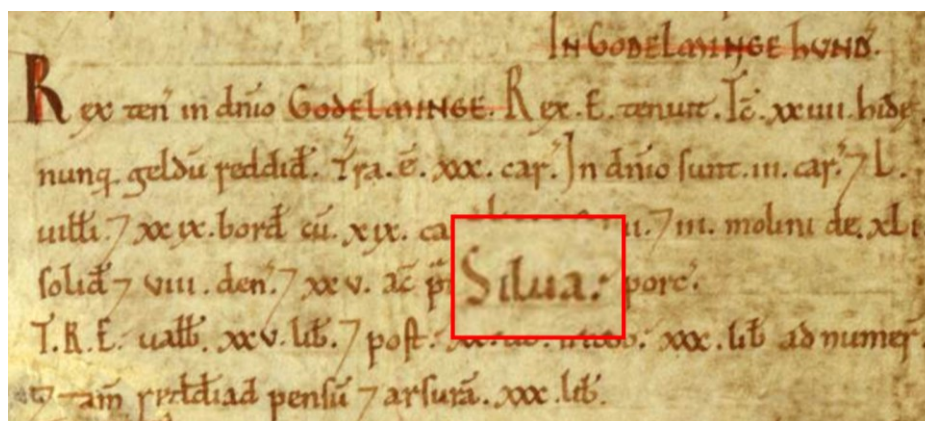


Figure 2: GDB 30v, Godalming, the woodland formula.

### Formatting of the text

The expansion of the Latin, then, not only makes the text easier to read but also extends its interpretative possibilities. So too do graphical forms, design elements if you like, that are used to convey meaning. The forms of initial letters and place-names, for example, can indicate the status of the land described. Thus, an initial square I and rubricated uppercase place-name, that is a place-name ruled though in red, can indicate an estate centre, a manor. By contrast, a rustic I and lower case unrubricated place-name point to a subordinate holding, either a berewick or sokeland (figure 3). The forms reveal estate structures that are otherwise invisible in the text. In the Cambridgeshire hundred of Papworth, for example, three extensive estates can be identified in this way, the interlocking of their elements indicting they were of some antiquity in 1086 (figure 4).<sup>5</sup>

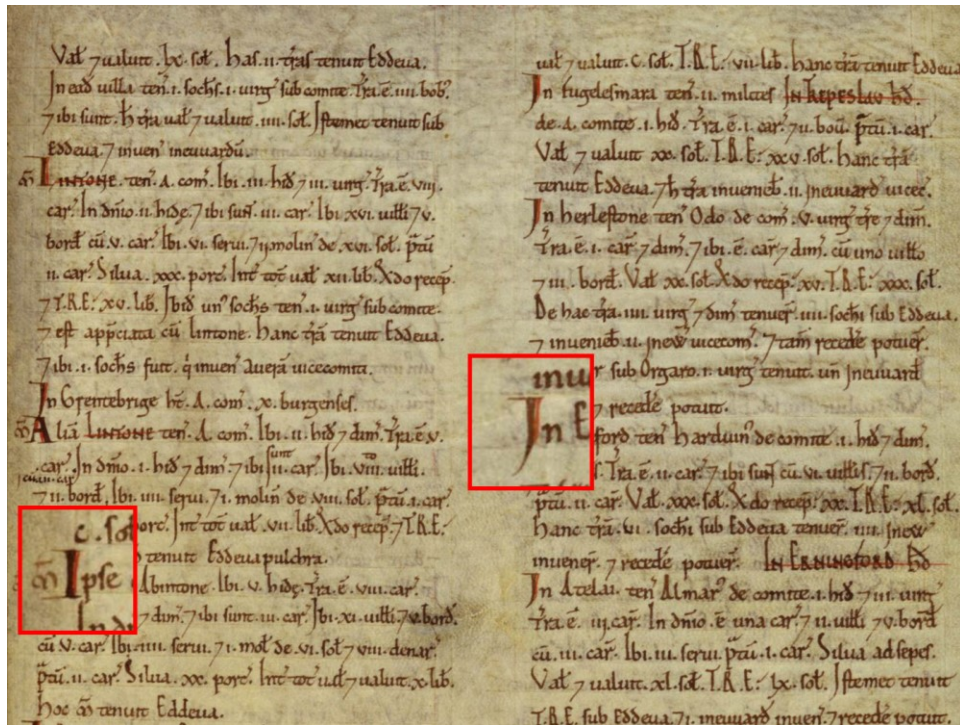


Figure 3: GDB 194, square and rustic initial letters. The occasional change of one form to another indicates that the devices were used consciously to convey meaning.

<sup>5</sup> David Roffe, 'McLuhan Meets the Master: Scribal Devices in Great Domesday Book', in *Domesday Now*, 81-108.

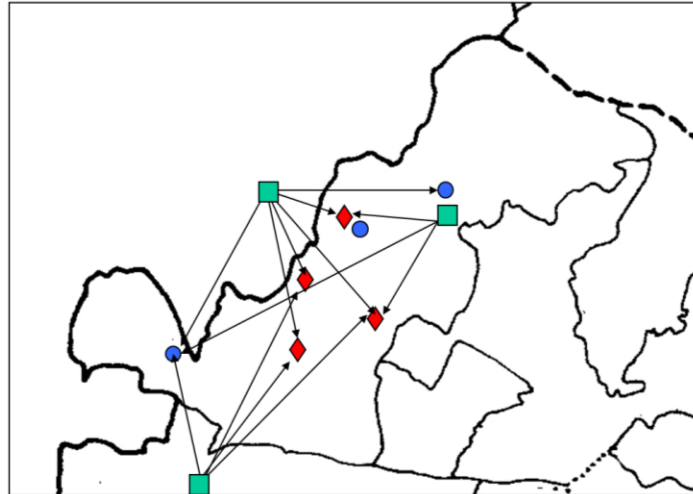


Figure 4: *estate structures in Papworth hundred, Cambridgeshire as reconstructed from initial letter forms. Square=manorial entry, circle=berewick, diamond=freeholding.*

Equally important is the general layout of the text. Blank lines can indicate different sections that are not otherwise signalled. In Kent, for example, they distinguish blocks of land in the same hundred, while in Lincolnshire they are sometimes intimately related to the tenant-in-chief's title, that is, they delineate pre-Conquest patterns of tenure and lordship. Schedules, that is discrete lists of lands or tenants incorporated wholesale into the text, can also be detected in the same way. Those attached to the king's lands in Staffordshire, for example, indicate that the scribe had lists of waste lands at his disposal. Finally, the scribe might signal exceptional material by the use of paragraphoi, gallows signs, that are added to the margin to bring the entries to the reader's attention. They are most regularly found in the account of boroughs where they signal non-customary lands, that is land that owed dues to a lord other than the king.<sup>6</sup>

### **Additions and corrections**

GDB is predominantly the work of one scribe. He was not a simple copyist, but compiled his account from various sources. Stints, that is writing sessions, can sometimes be detected from changes in the hand and pen. The scribe also reviewed what he had written and added to his text and corrected it in various campaigns. So, he might amend or gloss – insert explanatory material – as he wrote in the initial drafting. This is characterized as 'current' proofing work. Then there are changes that postdate initial drafting but precedes the rubrication of the text. Rubrication followed as a third stage and minor additions and corrections were subsequently added by a second scribe known as scribe B. Finally, the main scribe might return to his work after rubrication or even after B had finished his work.

The text, then, often has a complex stratigraphy. Mapping it is central to an understanding of the scribe's sources, his working methods, and the order of writing. That understanding remains incomplete. There is no consensus on the immediate sources of the text: for some it is was primarily based on circuit returns, that is drafts that

<sup>6</sup> *Ibid.*

encompassed a number of counties, for others the scribe drew on diverse sources, notably geographically arranged returns. In consequence, there is no agreement on scribal methods. Only in the order of writing is there some degree of common ground and there the details remain to be clarified. The new edition will provide the data on which consensus can be built.

### **The significance of textual forms**

All that is a preliminary to an appreciation of the scribe's developing programme. GDB was not conceived fully formed. When the scribe started with the Yorkshire folios, he had a very imperfect grasp of the nature of English society and how to represent it. A particular problem was how to deal with free men and sokemen: should they be categorized as lords or tenants? Many other issues also had to be sorted out. An understanding only gradually emerged of what was required and what was sufficient for the task in hand. It was not until about a third of the way through that the scribe settled down to a more or less standard form. In consequence the import of key concepts such as the manor changed in the course of writing.<sup>7</sup>

It follows that the meaning and significance of the information recorded in GDB depends not only on how it is recorded but also on when. Much of our understanding of regional variations in society and the economy have been based on differences in expression in GDB. It is now clear that many such are essentially Domesday artefacts. The distribution of free holds, for example, is skewed by the scribe's waning interest in the class as he wrote. The new edition provides the information on which a reader can base a judgement at any particular point in the text.

### **Editorial additions**

The Farley text has no editorial additions apart from folio numbers. The new edition additionally supplies a number for the quire, that is the gatherings or pamphlets that make up the MS. The unique entry references of the Phillimore translation are also given to enable cross reference to other Domesday datasets and resources. None of these data are interpretative. By contrast, the identification of the hundred in which each parcel of land was situated is derived from analysis and later sources. It is, however, important for appreciating hundredal structures and sources behind the text. I am using Frank Thorn's analyses and maps that are part of the Alecto translation.

### **Encoding and publishing the text**

Printing such a massive work would of course be prohibitively expensive. Could we expect large sales for a Latin text, even with the new translation that is planned? Well, no. And, moreover, it would not be possible to represent typographically all of the information that is recorded. How, for example, would the six periods of addition and correction be represented? So, I shall be publishing the text, on open access, on the internet.

To do so, I am marking it up in eXtensible Markup Language more or less in the form recommended by the Text Encoding Initiative. XML is a digital technology for encoding documents in a form that is machine readable but also comprehensible by mere mortals. TEI, provides a standardized framework for representing historical texts amongst

---

<sup>7</sup> David Roffe, *Decoding Domesday*, Woodbridge 2007, 176-82.

others in this form (figure 5).<sup>8</sup> The resulting code can be read directly by browsers, but they will only render the content as simple text. To bring out all the characteristics yet another technology is required. XSTL – eXtensible Stylesheet Transformation Language – is used to transform the code in the way desired.<sup>9</sup> It in effect tells the browser how to represent all the information that has been encoded and arrange the text accordingly. It will be possible, for example, to colour code different types of entry, recover all entries that were added at any one period, or even re-arrange the whole text by hundred.

```

<div type="fief" xml:id="CAM 31"><line>cadd hand="1" place="ml" period="3">.XXIX.</add>
<cadd hand="1" place="il" period="3">TERRA WIDONIS DE RAINVEDEVRT.</add></line>
<lb>
<div type="entry" xml:id="CAM 31-1"><hundred n="CAM-Armingford"><initial type="0">W</initial>IDO de
Rainbuedcort tencabbr type="1">et</abbr> In <place type="0"><span>Meirede</span></place> .iiii.
hid<abbr type="0">as</abbr>
<lb>
<abbr type="7">et</abbr> i. uirg<abbr type="1">atam</abbr> de rege. I<abbr type="1">er</abbr>ra
.<abbr type="0">et</abbr> .v. car<abbr type="1">ucis</abbr>. <abbr type="7">et</abbr> ibi sunt. In
d<abbr type="0">om</abbr>io <lb> dim<abbr type="1">idia</abbr> hida. <abbr type="7">et</abbr> ibi
.<abbr type="0">et</abbr> .i. car<abbr type="0">uca</abbr>. <abbr type="7">et</abbr> xv. bord<abbr
type="0">arii</abbr> cuc<abbr type="0">mc</abbr> .iiii. cot<abbr type="1">ariis</abbr> h<abbr type="0">
>abe</abbr>nt <lb> .iiii. car<abbr type="0">ucas</abbr>. Ibi un<abbr type="9">us</abbr> seruus.
<abbr type="7">et</abbr> ii<sup></sup>. molini de .x. sol<abbr type="0">idos</abbr>
<abbr type="7">et</abbr> viii. den<abbr type="1">arios</abbr>. <lb> P<abbr type="11"
>ra</abbr>t<abbr type="0">mc</abbr> .v. car<abbr type="1">ucis</abbr>. Past<abbr type="3">ur</abbr>a
ad pecun<abbr type="1">iam</abbr>. Int<abbr type="1">er</abbr> tot<abbr type="0">um</abbr> ual<abbr
type="0">et</abbr> lxx. sol<abbr type="0">idos</abbr>. <lb>
<abbr type="7">et</abbr> t<abbr type="1">a</abbr>nt<abbr type="a">un</abbr>d<abbr type="0">em</abbr>
q<abbr type="2">uan</abbr>do recep<abbr type="0">it</abbr>. T. R. E<abbr type="15">:</abbr> c.
sol<abbr type="0">idos</abbr>. Hanc t<abbr type="1">er</abbr>ra<abbr type="0">mc</abbr> tenuerunt <lb>
<hr/><folio no="200a"/><xv<subst>del hand="2" type="ow" period="2">. </del><cadd hand="2" place="11"
period="2">1.</add></subst> soch<abbr type="0">emann</abbr>1. De his .x. habuer<abbr type="0">
>unt</abbr> .ii. hid<abbr type="0">as</abbr>
<abbr type="7">et</abbr> dim<abbr type="1">idiam</abbr> uirg<abbr type="1">atam</abbr> de soca <lb>
S<abbr type="0">anctae</abbr> Edeldride de Ely. quo<abbr type="4">rum</abbr> un<abbr type="9"
>us</abbr> t<abbr type="1">er</abbr>ra<abbr type="0">mc</abbr> sua<abbr type="0">mc</abbr> nec dare

```

Figure 5: xml mark-up of part of GDB 200b.

For the first time it will be possible to search the text electronically, for both content and form. In the last twenty years the study of the diplomatic of the text – the form of words employed – has emerged as one of the most powerful approaches to Domesday Book. It has elucidated the nature of tenure and the referents of various terms among many other matters. Often, however, the analysis has been based on samples of the text rather than the whole because of the sheer volume of data. With the new edition it will be possible to use the entire dataset. Moreover, searching will not be confined to content. It will also be possible to search for forms, whether it be of initial letters, place-names, or whatever.

## Conclusion

This is a long term project. So far I have completed seven out of the thirty-one counties described in GDB. There's still a long way to go, but I hope you agree that the potential for the project is awesome. Once we have a complete text that is not only machine-readable, but also encodes so much of the information that a critical reading of the text affords, analyses will be facilitated that have hitherto been difficult or impossible. And, yes, for those of you who balk at Latin, there will be a new translation that takes into account the new discoveries.

©David Roffe, 2018

<sup>8</sup> <http://www.tei-c.org/index.xml>.

<sup>9</sup> <https://en.wikipedia.org/wiki/XSLT>.