THE DEEDS PROJECT:
TOWARDS THE CONTENT ANALYSIS OF ENGLISH PRIVATE CHARTERS OF THE TWELFTH AND THIRTEENTH CENTURIES

Introduction

The Norman contribution to the administrative history of medieval England is a commonplace of academic inquiry into the period. Among the many novelties which followed the Conquest of 1066, there is one in particular whose consequences have blurred our understanding of social, political and economic change for the nearly two and a half centuries separating the Conquest from the end of the reign of Edward I in 1307. That particularity is the custom of not including a date of issue in records recording property transfer. Such records, known as deeds or charters, were the most ubiquitous records of the time and it is estimated that for the twelfth and thirteenth centuries alone over a million have survived to this day as originals or cartulary copies.

There was considerable diversity on the Continent at the time of the first millennium in the tendency to include, or exclude, a date reference when property conveyances were recorded in writing, depending upon where and when they were issued. There can be no doubt that the closer the tie between the issuer and the Roman tradition, the greater the likelihood that a charter would bear a date. There is increasing evidence to suggest that in Burgundy, as in Normandy, the custom of including dates declined steadily from the late tenth century until the middle of the eleventh century, when they became quite rare. Although situated within the Frankish realm, these territories lay well outside the formal control of the French king. It is above all this detachment from royal administrative supervision which may explain

why, when the charters of the French monarchy continued to be dated, those of some, at least, of the more independent outlying provinces did not. In this regard it was England's fate that the Norman Conquest took place when it did, for with William I came the then-current Norman custom of not dating charters at any level of society. The Norman dukes, with an administrative system developed when ties with the French Crown were weak, were clearly unconcerned about the formal letter-writing conventions which European monarchs had adopted from the papal chancery. They bothered little with traditional formality, dispensed with all that was not absolutely essential to the message they wished to convey, and only occasionally included a date. They sought brevity and conciseness in their charters and carried that administrative principle with them to England where it became a long-standing tradition in its own right. It was undoubtedly this tradition, rather than a conscious desire to dispense with the concept of time, that led to the enduring English phenomenon of the undated charter. In Normandy, dating returned to the charter text as royal administrative authority spread in the region after 1204, while in England it was with the accession in 1189 of Richard I, whose administrative experience was rather Continental than English, that dating was regularly introduced to records emanating from his royal chancery. Attachment to tradition being what it is, it still took well over a century for that custom to be adopted consistently by those who drafted charters elsewhere in the realm. It was not until the early years of the reign of Edward II (1307-27) that it became customary to include dates in private conveyances.

The primary objective of the DEEDS research project is to develop a methodology for the analysis of the medieval "letter" or charter. The core of our approach is that unknown attributes of any given charter can be determined by comparison with a set of similar charters whose attributes are known. Once the

methodology has been established, it can be used to make chronological evaluations, to identify related forgeries, to distinguish between and/or identify authorship; and to undertake a wide range of historical research based upon quantitative analyses of content. The development of a methodology to date undated medieval charters has been the primary occupation of the DEEDS Project over the past eight years. The traditional method of ascertaining at least a relative chronology for undated records has been through the association of personal names in those records with their counterparts in dated sources. This approach works reasonably well when one, but preferably two or more office-holders whose dates of tenure are known appear in the same record. It becomes distinctly hazardous, however, in the case of individuals to whom a clearly defined chronology cannot be attached, since namesakes, even when occurring in groups (as they do from one generation to the next in rural environments), can often be shown to be different people. And even when the identification is certain there is usually no way of telling at what point in the life-span of that individual a given record was issued. For lack of sufficient contextual evidence inaccurate attributions can easily be made in the case of unknown individuals and dubious dates consequently assigned to the records in which they appear. Furthermore, once a circa date has been attributed to such a record and the work goes into print, that date tends thereafter to take on a historical value of its own. It is given the benefit of the doubt by subsequent scholars as though it actually belonged to the record in question. In this manner, dates are assigned to undated charters which in turn are used as evidence in the dating of other undated charters. The potential for error in this process is rather greater than one would like it to be, and probably far greater than one could imagine. In a recently published posthumous article by John Horace Round, the author points out how the respected editor of the

Colchester Cartulary misdated a dated royal charter by more than 200 years because he did not know that the king's style as registered in the dating clause, "Henricus Rex Anglie et Francie et Dominus Hibernie", designated Henry VI rather than Henry III[1] When errors of this sort can be made with dated royal charters, how much more vulnerable might be dates assigned through the association of personal names to undated private ones.

Problems arising from the misidentification of different individuals bearing the same name, or of the same individual bearing the same name, not to mention the same individual bearing different names, are compounded in those records whose witness lists were omitted, as often happened when they were copied into a cartulary. In the case of records in which no more than two or three otherwise unknown names appeared, other means were obviously needed to identify chronological change. Paleography was not an option as the great majority of the sources from the period have survived only in later copies. For the same reason, sigillography was also insufficient, regardless of its importance in those relatively rare cases where originals exist together with their seals.

Seeking an alternative solution, DEEDS turned to a route suggested by the English medieval historian, Sir Frank Stenton, eighty years ago. Faced with the reality that only five percent of the records he was editing bore dates, or could be assigned accurate dates based upon recognizable names or events cited in them, he became convinced after making a meticulous examination of his sources that charter chronology was inextricably tied to the growth and development of the formulae which appeared in them.[2] The term "formula" is somewhat ambiguous as it suggests a set of words whose order does not change regardless of context. Since nothing could be further from the truth in terms of the medieval legal phraseology of the

twelfth and thirteenth centuries, DEEDS prefers the expression “word-pattern” to “formula” and uses it to designate any group of two or more words which occurs in more than one record.

In general, any charter is an official legal document written or issued by a religious, lay or royal institution and therefore can be treated both as an independent entity and as the object of modelling and analyses. It is important to recognize that history reflects the individual circumstances which determined changes in the construction of legal texts and the constant adoption, formulation and adaptation of words and word expressions. Underlying this whole process of change, however, is the phenomenon, common to all languages at all times, of obsolescence and word immediate reflection of social change.

In 1983, using what was then considered “new” methodology, the DEEDS Project produced a microfiche concordance of the content of the 957 charters in the secunda camera of the English Hospitaller cartulary of 1442[3]. Careful perusal of the concordance made it unquestionably clear that these charters alone contained tens of thousands of identifiable word-patterns, that is groups of two or more words in one record which were exactly similar to a corresponding group in another record. The wide variety of dates associated with the different patterns led to the supposition that rather than being predominantly static, with the occasional change resulting from a historical event or made in response to a royal statute, the language of the medieval charter was in constant flux. Taken as a whole, it seemed reasonable that if the chronology of change were clear, it should be possible to use the rates of appearance of word-patterns in the record, and their disappearance from it, to determine a fairly accurate date for any given record of this sort. In other words, chronological

groupings reflected in the word-patterns from the concordance suggested word-pattern matching to be a potentially viable means for dating the undated charter.

We believe (and preliminary analyses confirm) that chronological boundaries can be derived by studying the frequency distribution of terms and expressions extracted from a given charter over time. After taking into consideration any additional information that might be found in the text, such as the nature of the content, the structural organization of the text, topographical features, references to names of people and institutions, etc., final conclusions can be drawn by combining the results collected from the above analyses of each such component.

DEEDS Corpus of Medieval Charters

The DEEDS Project's Corpus presently includes more than 8000 medieval Latin charters from twelfth- and thirteenth-century England derived from printed sources. All charters included in the Corpus are dated internally or by the editor of the collection using internal evidence. Most of the charters chosen for our corpus are internally dated, usually by anno domini or regnal year, or both. Occasionally, papal or episcopal regnal years are cited, usually in episcopal charters. Sometimes, an historical event provides a date. To these, we have added undated charters which can be dated to a given year by internal evidence, such as the many published by R.E. Salter for the Oxford Record Society which contain the names of office holders who only held their post for a year at a time. Charters with "assigned" dates represent thirty-two percent of the total number of charters in the DEEDS corpus.

Obviously, the number of charters available for each chronological time span varies (Fig. 1), as does the accuracy of the chronological evidence which varies from the exact day, month and year to a range of several years. An average charter contains

about 200 words (1300 characters) its size is not specifically related to any time span or to the nature of its content. Very short or very long ones are relatively uncommon (Fig.2). Ninety-two percent of the charters in the corpus are represented by sixteen types, while five of these types alone make up nearly eighty percent of the total (Fig. 3). Those five are Grants (33%), Agreements (17.5%), Final Concords (12.4%), Confirmations (9.4%) and Quitclaims (8.8%).

One of the major challenges faced by the researcher working from published records rather than from manuscript sources is dealing with a myriad of editorial methods and principles devised by individual editors. Wherever possible, the computer is used to eliminate or circumvent such differences. Letter “v” is invariably read as “u”, and “j” as “i”; letters “c” and “t” are standardised in certain combinations. No distinction is made between upper and lower case letters, and punctuation is removed. More problematic are cases where editors (and sometimes medieval copyists as well) have abbreviated texts or replaced what they interpret as purely formulaic phrases with “etc.” We mark up obvious lacunae and retain abbreviations, but disregard them in research queries, and always retain the “etc.”

Source Document (SD)

SD (Fig. 4) is an XML document composed of the original text extracted from the printed source by scanning text from it, converting it to electronic format using Optical Character Recognition (OCR) and complementary information derived 1) from the charter itself, 2) from the corresponding edition of the cartulary or collection of charters from which it is taken, or 3) from external sources. SD is a plain text document defined and written in XML. SD can easily be parsed by a computer or read by a person. We developed a proprietary programme called

“Document Manager” (DM) to facilitate the creation and maintenance of SDs. SDs are stored in the DEEDS Project Database allowing for content indexing and convenient access to the information over the Internet.

The purpose of introducing SD is to encapsulate information and establish logical links between the text of the charter and complementary information related to the charter in sharable and portable form.

The use of XML has made it possible to implement a highly flexible repository of textual data, along with related auxiliary information. Each SD and its key data (a combination of the arbitrary code for the cartulary in which it was published and the number if the charter inside the Cartulary) occupies a single row of the database table. The use of a database gives access to the indexing and exploratory facilities of modern database management systems and ensures the consistency and integrity of the information. It should also be noted that the hierarchy used in SD could also be implemented using a "pure" database solution. This database solution would require the creation of multiple database tables and the setting up of a complex of links between them. The administration, maintenance and application programming of the "pure" database solution would involve much more effort than an XML based solution. Therefore, the use of XML was a logical choice for developing the DEEDS Corpus of Source Documents. An additional benefit of the use of XML is the widespread availability of XML Parsers helping to reduce significantly the complexity of application programming.

SD is composed of the following four main elements: **content**, **data**, **notes** and **markup** where the last three elements hold complimentary information which is linked to the given charter.

- *"Content*". This element stores the original text of the charter as it appears in the printed source. It has no child nodes, therefore the integrity of the charter's text is fully preserved.
- *"Data*". This element stores information regarding properties attributed to a given charter, which are useful for us in dating. Wherever possible, we attribute the following information to each entry in the corpus: 1) the type of date associated with the charter (assigned by an editor or internal; if internal, by Anno Domini, "regnal" year, "feastday", or event), 2) the type of legal transaction it records (grant, agreement, confirmation, quitclaim etc.); 3) whether the source is an original record or a copy and, if a copy, the nature and date of the source; 4) the place of issue, 5) the parish and county concerned, 6) the general nature (religious or lay) of the issuer and recipient, 7) the specific nature of the issuer and recipient (religious order or institution, or an individual's name, title and occupation); 8) in the case of religious houses, the name of the house and the dates of foundation and dissolution;
- The "*Notes*" element stores editorial information from external sources which is related to the given charter, and divided by theme.
- The "*Markup*" element implements the "Virtual" Markup by storing the addresses of all markup tags. Virtual Markup was introduced to establish an indirect system of textual markup using XML tags. These tags are mapped to the charter's text rather then being physically inserted into it.

Document Manager

Document Manager (Fig. 5) is a software program developed by the DEEDS Project to facilitate the addition to, modification of, and removal of charters from the DEEDS corpus. It is implemented as a Windows Form composed of data fields grouped into a series of data entry screens. The "Save" operation can be initiated as soon as appropriate data has been entered into the form fields and will result in the generation of the XML format of the SD. This format would replace, or be added to, the database housing the DEEDS corpus. Any newly created or modified SD is in full compliance with the SD Document Type Definition (DTD) (Fig. 6) and with related validation rules. This program eliminates all errors caused by mis-spellings and typos in XML tags and attributes. DM also offers convenient categorized access to the charter itself and to complimentary data.

In addition to the above-mentioned functionality, the program allows full access to such auxiliary data as our Lists of Cartularies, Ecclesiastic institutions, and Topographical locators (such as country, county or parish). All lists are XML documents stored in the database, also created and managed using DM.

Virtual Markup

Virtual Markup represents that part of the unique technology developed by the DEEDS Project to perform content-oriented quantitative analyses of medieval charters. Charter content may be described and modeled in different terms depending on the goal of the research being undertaken. It could be categorized in terms of formulae (i.e. diplomatic parts), content roles (i.e. subjects of legal action, references to chronological information, people's names, descriptions of properties, etc.) or a literal phrase.

The *Markup* element is implemented through a collection of “Layers”. Each layer is tightly linked to a specific category. The Layer itself is a collection of items, in which each item carries the precise address of the markup tags associated with a given expression found in the text. The **Item** stores locations of the start and the end of the given phrase relative to the beginning of the text. There is no limitation to the number of Markup layers or items which may be introduced. In practice Virtual Markup is created, maintained and visualized by the Document Manager.

Conventional markup requires markup tags to be embedded into the text. Virtual Markup is a generic alternative to it and requires that only the addresses of markup tags be collected and saved as child nodes of the XML document "tree", while the text of the charter remains in its original state. Here are some benefits of the use of Virtual Markup:

- It imposes no mandatory rules for the layout of the charter and may therefore be used with any charter type, including ones composed of multiple documents (as sometimes occurs in "Confirmations").
- Ability to build and maintain an unlimited number of independent categorizations for a given charter.
- Markup nodes of independent categories can overlap, a possibility that is completely excluded in the case of conventional XML markup.
- Fast and reliable access is provided to marked-up information, and the text of a charter, without the need to resort to the use of resource-consuming navigation over all branches of XML tree queries

The presence of categorical markup, along with other information attributed to the documents in the DEEDS Corpus, allows us to produce, on demand, filtered versions of the Corpus (so-called “projections”). A projection of the Corpus might

include only the documents fitting the criteria of certain selected attributes and, for those, the original text could be replaced by the content of its marked-up segments.

Editors do not concur in their assessment of record types. As a consequence, DEEDS attaches its own designation in a standardised approach based on words of disposition (GRANTS invariably include forms of dedi and/or concessi, sometimes together with liveravi, tradidi, dimisi, assignavi, contuli, legavi, and donavi; CONFIRMATIONS incorporate forms of confirmavi, and concessi et ratam habui; QUITCLAIMS, forms of quietum clamavi, sometimes in conjunction with remisi, concessi, recognovi, resignavi, relaxavi, confirmavi, and dimisi; LEASES, forms of dedi, dimisi, concessi, and tradidi; SALES, forms of vendidi, quietum clamavi, concessi, forisaffidavi, and confirmavi; AGREEMENTS, forms of convenit, or conventio and compositio). In order to confirm the totality of dispositive words and phrases concomitant with record types, those words and phrases are marked up and lists of patterns established for each type. In addition to the words of disposition the standard charter is a composition of PROTOCOL, CORPUS, and ESCHATOCOL is further broken down such that the parts of the PROTOCOL may be identified as the invocatio, intitulatio, and salutatio; the CORPUS according to the divisions of specific record types (generally the promulgatio or notification, the narratio or background, the dispositio or words of action, and the conditional clausulae); and the sections of the ESCHATOCOL as the sealing clause, the witness list, the place of issue and the date.

We use projections of the original charters in quantitative content analyses and chronological evaluations to eliminate the negative influence of editorial errors, spelling and formatting inconsistencies on the final estimate. The projection of the Corpus consists of documents in HTML format derived from XML formatted SDs by using Extensible Style Sheet Language (XSL) and proprietary programming routines.

Quantitative Content Analyses and Chronological Evaluations

Quantitative analyses and evaluations based on the examination of vocabulary involves extracting combinations of two or more adjacent words, so called **word-patterns,** in consecutive order from the text of the charter and then examining occurrences of each pattern in the Corpus. Lists of the occurrences of each word-pattern are then analysed to generate estimates of the unknown properties (attributes) of the given charter.

We analyze the distribution of frequencies of word-pattern occurrences (Analyses of Frequencies) to estimate the unknown attributes of a charter by comparing them with charters in the Corpus. The same approach taken with specific projections of the Corpus would provide more specific and focused results for the parameters selected.

Analyses of Frequency begin by extracting sets of two or more consecutive words from the text of the given charter. The program starts from the first word and then moves down to the end of the text, extracting a fixed number of words in one-word steps. The generation of sets ends when the word-pattern size reaches the size of the charter text or when no word-pattern in the given set occurs elsewhere in the Corpus.

When generation of the set has been completed, the program begins to search the Corpus for occurrences of each extracted word-pattern. The resulting distribution of occurrences is treated as an independent collection and is grouped along with the properties of the individual word-pattern. There can be wide variation in the number and length of word-patterns in each set and in matches (hits) with patterns in the Corpus. After all word-patterns in the set have been processed, the system computes

an estimate and expected margin of error for each pattern. In the following stages, those word-pattern estimates are scaled, weighted and then grouped together into the word-pattern set-level estimate and, again, the **central tendency estimate** and its margin of error are computed at that level. Further set-level estimates are scaled, weighted and then merged into the charter level. Computing an estimate and margin of error at the charter level concludes the Analyses of Frequencies for the given **projection**. Weighting factors used in the computations are always computed dynamically and depend on the individual properties of a given **distribution**. The goal of dynamic weighting is to reduce distortion caused by natural variations in word-pattern sizes and properties between layers on all levels. Eventually, the final estimate of unknown attributes may be calculated by merging together estimates produced by the processing of multiple projections within the same charter.

Our methodology for chronological evaluation is based on the premise that one can make accurate judgments about the chronology of a charter by comparing it with other charters in the Corpus. Therefore, a match between the known attributes of a given charter with corresponding attributes in the group of charters allows us to assign values of other attributes back from the group to the given charter. Any feature of the charter that can be derived directly or indirectly from the text, reliably measured and then matched to the same features of other charters in the Corpus, can be used as an attribute. Because our Corpus is composed of independent documents, the result produced by any matching of attributes can be nothing but a collection of attribute values. An estimate of the attribute value associated with each collection can be obtained by using statistical techniques for measuring the central tendency of distribution, together with an estimate of the expected margin of error. In subsequent stages of processing, results produced by any similar collection are gathered in groups

and each group is again treated as an independent collection. Eventually, a final estimate of the given attribute is computed by processing the last collection generated, which itself represents the result of multiple mergers of collections produced during the intermediate stages of processing.

Fig. 7 displays a system output produced by using the entire Corpus (at the top) and the content type based Corpus Projection (at the bottom).

Summary of Research Intent and Results

The primary objective of the DEEDS research project is to develop a methodology for the analysis of the medieval "letter" or charter. Once established, that model can be used: to make chronological evaluations, to identify forgeries, to distinguish between or identify authors; and to undertake a wide range of historical research based upon analyses of content.

Preliminary analyses confirm that chronological boundaries can be assigned to undated charters by comparing the frequency distribution of terms and expressions generated from a given undated charter against an existing set of records whose attributes have known values. Hand in hand with the establishment of correct chronology is the identification of forgeries, since the purpose of forgery is to change past intent. DEEDS proposes to approach the identification of forgeries through the analysis of content, particularly in terms of words in context, their placement in the text, and the frequency of their usage. Sections containing words, phrases or concepts which appear to be outside their normal chronological context will be singled out as having possibly been entered by an "improver" or forger. Few forgeries entail complete records; it is the component parts which are likely to have been tampered with.

By adding chronological, spatial, lexical and structural components to inquiries of this sort, researchers at DEEDS anticipate that it will be possible to improve existing methods of authorship evaluation. The result in the case of charters will be to identify the work of individual scribes, the date range associated with their activity, and their patron or patrons.

The DEEDS Project uses XML for the processing and analyses of Latin Medieval Charters. We found XML to be a reliable and universal medium conveniently allowing us to combine and manipulate the original textual material and the attributed data. Today's progress in XML processing technology provides researchers with additional powerful technologies, such as XML Schema, as real alternatives to traditional DTD and data transfers between XML and the database to which it is applied as a standard feature of all leading Database Management Systems.

Michael Gervers, PhD,
Director, DEEDS Project,
Professor of History,
University of Toronto
e-mail: 102063.2152@compuserve.com

Michael Margolin, BA,
Programmer / Analyst, DEEDS Project,
University of Toronto
e-mail: m.margolin@utoronto.ca

Appendix 1

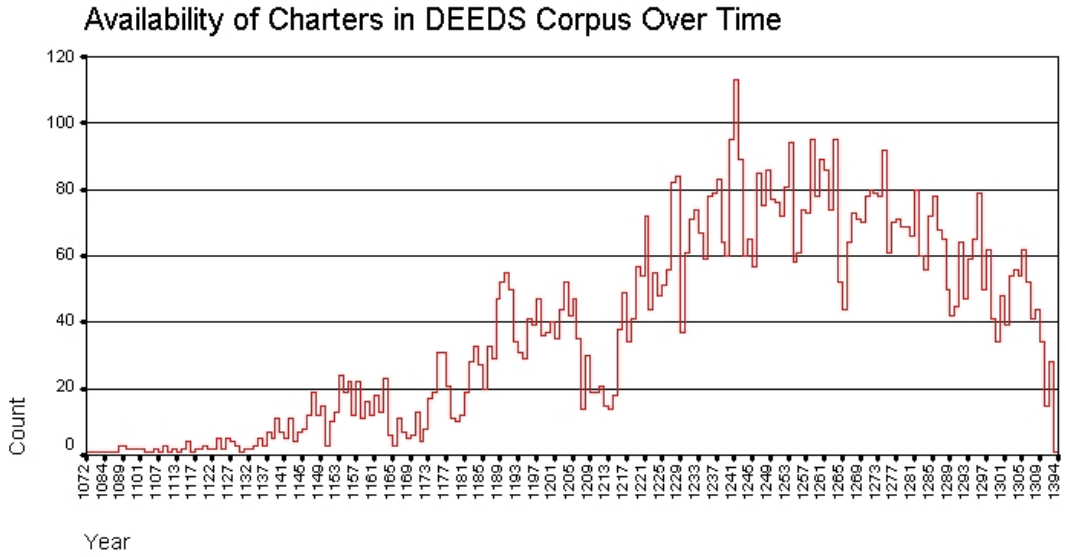


Figure 1 Availability of Charters in the DEEDS Corpus Over Time

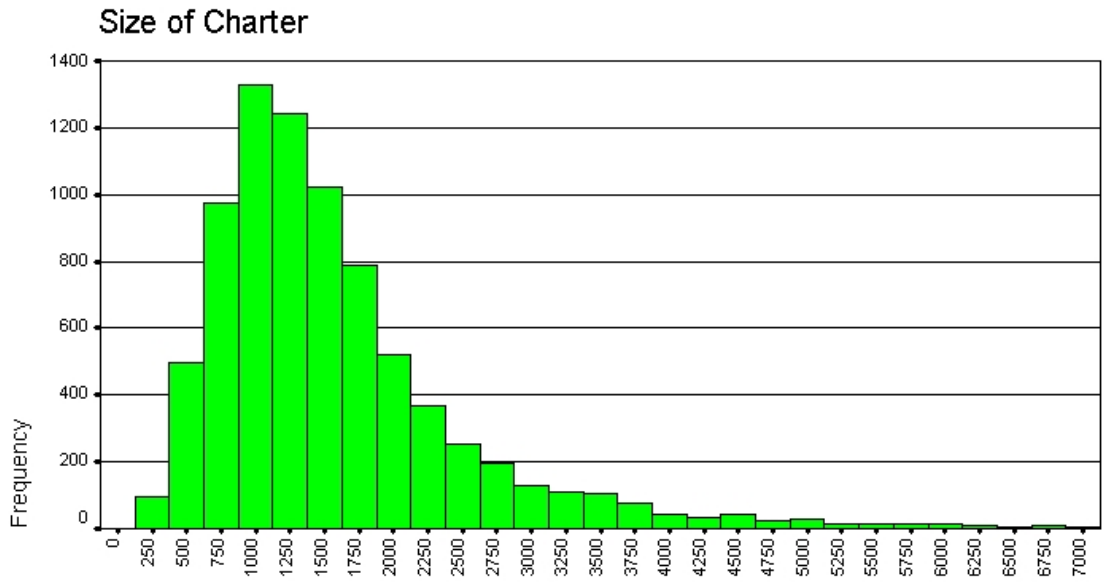


Figure 2 Variation of Charter Size in the DEEDS Corpus

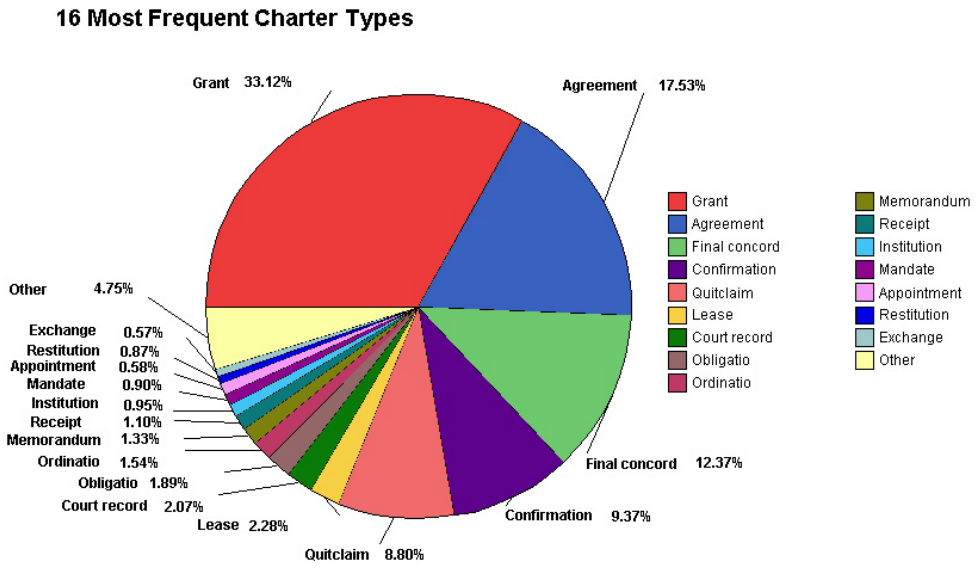


Figure 3 Most Frequent Charter Types

<sourceDocument dnum="00880049" created="2004-03-08" cartulary="Knights of St John secunda camera">
<content>Sciant presentes et futuri quod ego Walterus carpentarius filius Henrici molendinarii de Hengham Sibile concessi dedi et hac presenti carta mea confirmavi Rogero de Othulvesho clerico pro servicio suo et pro viginti solidis sterlingorum quos michi dedit in gersumam unam acram terre et tres rodas cum pertinentiis in dicta villa de Hengham Sibile iacentes ex opposito mesuagii Iordani carpentarii inter terram Edithe Smuches et terram que Galfridi Berde fuit et pasturam duarum acrarum terre iacencium propinquius mesuagio Godefridi coci apud le Motstowe ibidem singulis annis cum segetes amote fuerint scilicet de terra Iohannis filii Alani habend et tenend de me et heredibus meis predicto Rogero et heredibus suis et suis assignatis et quibuscumque vel quandocumque predictam terram et pasturam cum suis pertinentiis dare vendere et assignare voluerint et eorum heredibus libere quiete bene et in pace iure in feodo hereditarie et in perpetuum reddendo inde annuatim michi et heredibus meis vel meis assignatis ille et heredes sui vel sui assignati et eorum heredes 7d ad duos terminos anni videlicet ad Pascha 3d [et] obolum et ad festum sancti Michaelis 3d et obolum pro omnibus serviciis consuetudinibus exaccionibus sectis curiarum et universis secularibus demandis salvo tamen servicio domini regis scilicet ad scutagium quando evenerit ad plus et ad minus unum obolum Et ego predictus Walterus et heredes mei vel mei assignati warantizabimus defendemus et acquietabimus totam predictam terram et pasturam cum omnibus suis pertinentiis predicto Rogero et heredibus suis vel suis assignatis et eorum heredibus per predictum liberum servicium contra omnes gentes in perpetuum In cuius rei testimonium huic presenti scripto sigillum meum apposui Hiis testibus Adam Gravassall Iohanne fratre suo Willelmo filio Willelmi Willelmo le Bloy Martino crok(er) Roberto hostage Roberto Fovel Huberto clerico Huberto crok(er) Martino le Boll Waltero de Hauekwode Iohanne de le Hill clerico et aliis Datum apud Hengham Castrum die sancti Dionisii anno regni regis Henrici filii regis Iohannis quadragesimo sexto</content>
<data>
<general>
<origin>Lay</origin>
<transactionType>GRANT</transactionType>
<copySource>CartularyCopy</copySource>
<status ref="">Primary</status>
</general>
<dateInfo>
<singleDate>1262-10-09</singleDate>
<lowDate>0000-00-00</lowDate>
<highDate>0000-00-00</highDate>
<main>Internal</main>
<details>Feast</details>

```
            <details>Regnal</details>
        </dateInfo>
        <locationInfo>
            <location type="issued">
                <locSet>
                    <place>HEDINGHAM, CASTLE</place>
                    <county>ESSEX</county>
                    <country>ENGLAND</country>
                </locSet>
            </location>
            <location type="property">
                <locSet>
                    <place>HEDINGHAM,SIBLE</place>
                    <county>ESSEX</county>
                    <country>ENGLAND</country>
                </locSet>
            </location>
        </locationInfo>
        <parties type="givers">
            <person type="LAY">
                <pname type="Individual" role="" link="">Walter the Carpenter son of Henry the miller of
Hedingham, Sible</pname>
            </person>
        </parties>
        <parties type="recievers">
            <person type="LAY">
                <pname type="Individual" role="" link="">Roger of Othulvesho, clerk</pname>
            </person>
        </parties>
    </data>
    <notes/>
    <markup>
        <layer type="numeral">
            <item start="1070" end="1072"/>
            <item start="1115" end="1129"/>
            <item start="1160" end="1172"/>
            <item start="1366" end="1377"/>
        </layer>
        <layer type="name">
            <item start="37" end="104"/>
            <item start="156" end="184"/>
            <item start="378" end="397"/>
            <item start="411" end="425"/>
            <item start="440" end="454"/>
            <item start="524" end="538"/>
            <item start="622" end="642"/>
            <item start="693" end="699"/>
            <item start="1395" end="1403"/>
            <item start="1554" end="1560"/>
            <item start="1766" end="1781"/>
            <item start="1782" end="1800"/>
            <item start="1801" end="1824"/>
            <item start="1825" end="1841"/>
            <item start="1842" end="1858"/>
            <item start="1859" end="1874"/>
            <item start="1875" end="1888"/>
            <item start="1889" end="1904"/>
            <item start="1905" end="1921"/>
            <item start="1922" end="1937"/>
            <item start="1938" end="1958"/>
            <item start="1959" end="1985"/>
        </layer>
        <layer type="location">
            <item start="333" end="347"/>
            <item start="544" end="555"/>
            <item start="2006" end="2021"/>
        </layer>
```

```
        <layer type="property">
            <item start="135" end="140"/>
            <item start="189" end="197"/>
            <item start="209" end="237"/>
            <item start="258" end="266"/>
            <item start="267" end="283"/>
            <item start="287" end="297"/>
            <item start="324" end="347"/>
            <item start="369" end="377"/>
            <item start="369" end="397"/>
            <item start="404" end="425"/>
            <item start="429" end="454"/>
            <item start="463" end="492"/>
            <item start="515" end="538"/>
            <item start="582" end="589"/>
            <item start="1070" end="1072"/>
            <item start="1108" end="1114"/>
            <item start="1115" end="1129"/>
            <item start="1136" end="1159"/>
            <item start="1160" end="1172"/>
            <item start="1185" end="1230"/>
            <item start="1286" end="1307"/>
            <item start="1320" end="1329"/>
            <item start="1371" end="1377"/>
            <item start="1696" end="1707"/>
            <item start="1730" end="1738"/>
        </layer>
        <layer type="date reference">
            <item start="2022" end="2106"/>
        </layer>
    </markup>
</sourceDocument>
```

Figure 4 Source Document (SD)

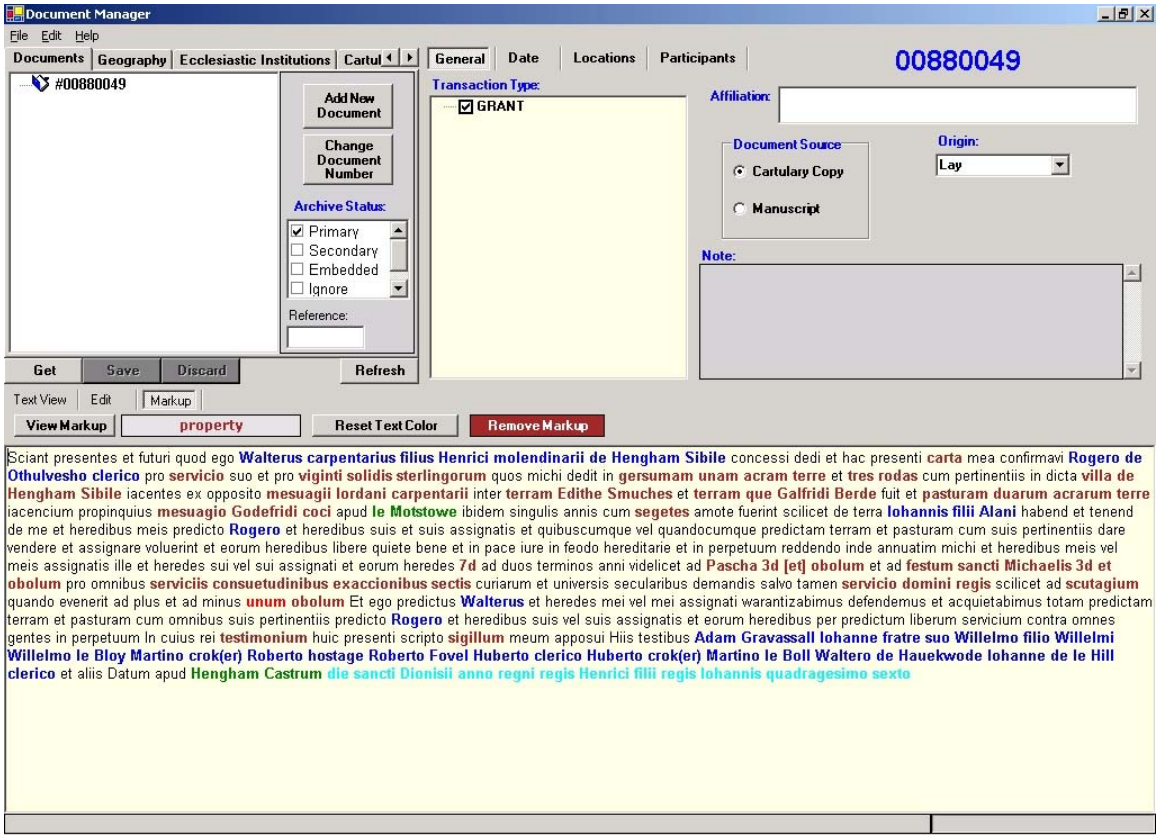
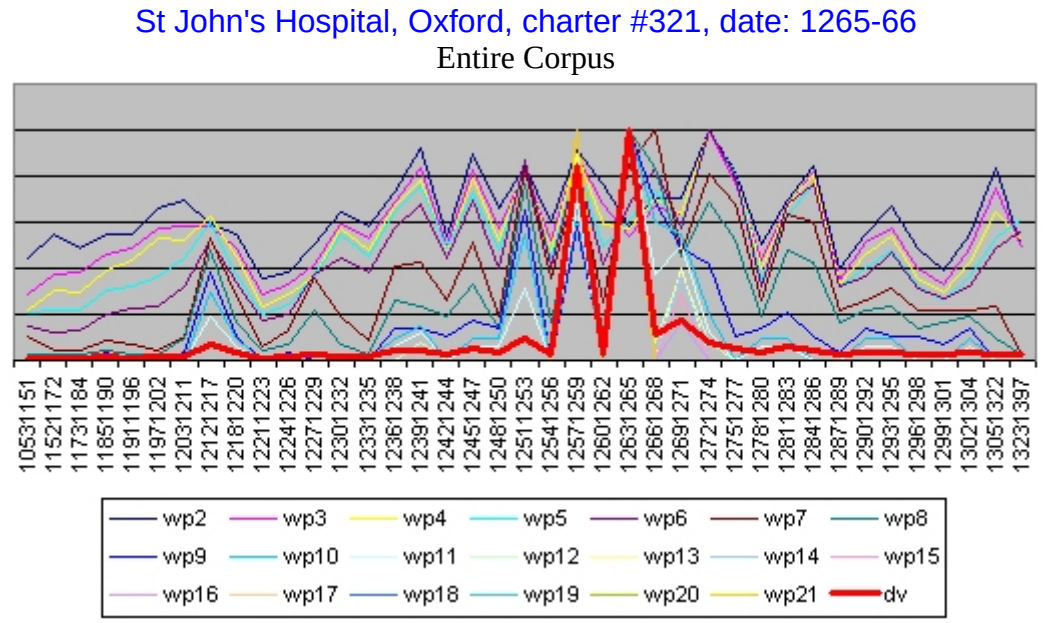


Figure 5 Document Manager

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v3.5 NT (http://www.xmlspy.com) by Michael Margolin (University of
Toronto/DEEDS Project) -->
<!ELEMENT sourceDocument (content, data, notes, markup?)>
<!ATTLIST sourceDocument
        dnum CDATA #REQUIRED
        cartulary CDATA #REQUIRED
        created CDATA #IMPLIED
>
<!-- text section -->
<!ELEMENT content (#PCDATA)>
<!-- data section -->
<!ELEMENT data (general, dateInfo, locationInfo, parties*)>
<!-- general -->
<!ELEMENT general (origin, transactionType+, status, affiliation?, copySource?)>
<!ELEMENT origin (#PCDATA)>
<!ELEMENT transactionType (#PCDATA)>
<!ELEMENT status (#PCDATA)>
<!ATTLIST status
        ref CDATA #IMPLIED
>
<!ELEMENT copySource (#PCDATA)>
<!ELEMENT affiliation (#PCDATA)>
<!-- date -->
<!ELEMENT dateInfo (singleDate, circa?, lowDate, highDate, dateSource?)>
<!ELEMENT singleDate (#PCDATA)>
<!ELEMENT lowDate (#PCDATA)>
<!ELEMENT highDate (#PCDATA)>
<!ELEMENT circa (#PCDATA)>
<!ELEMENT dateSource (main, details*)>
<!ELEMENT main (#PCDATA)>
<!ELEMENT details (#PCDATA)>
<!-- location -->
```

```
<!ELEMENT locationInfo (location*)>
<!ELEMENT location (locSet)>
<!ATTLIST location
        type CDATA #REQUIRED
>
<!ELEMENT locSet (place, county, country)>
<!ELEMENT place (#PCDATA)>
<!ELEMENT county (#PCDATA)>
<!ELEMENT country (#PCDATA)>
<!-- participants -->
<!ELEMENT parties (person*)>
<!ATTLIST parties
        type CDATA #REQUIRED
>
<!ELEMENT person (pname, title*)>
<!ATTLIST person
        type CDATA #REQUIRED
>
<!ELEMENT title (#PCDATA)>
<!ATTLIST title
        institution CDATA #REQUIRED
>
<!ELEMENT pname (#PCDATA)>
<!ATTLIST pname
        type CDATA #REQUIRED
        role CDATA #REQUIRED
        link CDATA #REQUIRED
>
<!-- notes -->
<!ELEMENT notes (note*)>
<!ELEMENT note (#PCDATA)>
<!ATTLIST note
        type CDATA #REQUIRED
>
<!-- markup -->
<!ELEMENT markup (layer*)>
<!ELEMENT layer (item*)>
<!ATTLIST layer
        type CDATA #REQUIRED
>
<!ELEMENT item EMPTY>
<!ATTLIST item
        start CDATA #REQUIRED
        end CDATA #REQUIRED
        type CDATA #IMPLIED
        num CDATA #IMPLIED
        date CDATA #IMPLIED
        datee CDATA #IMPLIED
>
```

Figure 5 Document Definition File (DTD)



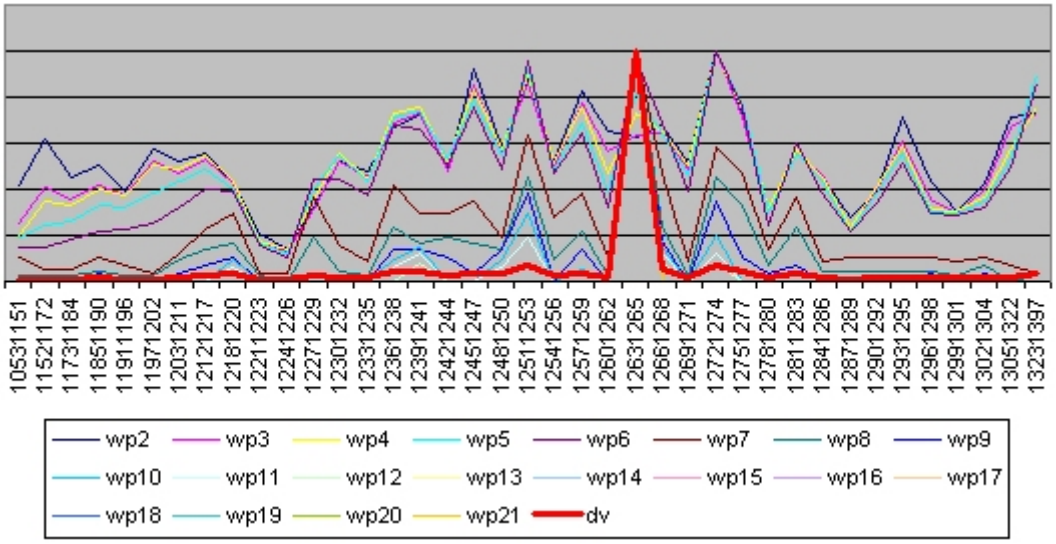


Figure 7 Chronological Evaluations

ADDITIONAL BIBLIOGRAPHICAL REFERENCES

CHIBNALL, Marjorie, “Dating the Charters of the Smaller Religious Houses in Suffolk in the Twelfth and Thirteenth Centuries”, in GERVERS, *Dating Undated Medieval Charters*, pp. 51-59.

DECLERCQ, Georges, “A New Method for the Dating and Identification of Forgeries? The DEEDS Methodology Applied to a Forged Charter of Count Robert I of Flanders for St. Peter's Abbey, Ghent”, in GERVERS, *Dating Undated Medieval Charters*, pp. 123-136.

GERVERS, Michael, “Changing forms of Hospitaller address in English private charters of the twelfth and thirteenth centuries”, in *The Crusades and the Military Orders: Expanding the Frontiers of Medieval Latin Christianity*, ed. Zsolt HUNYADI and József LASZLOVSZKY, Budapest: CEU, 2001, pp. 395-405

-------- AThe Dating of Medieval English Private Charters of the Twelfth and Thirteenth Centuries”, in *A Distinct Voice: Medieval Studies in Honor of Leonard E. Boyle, O.P.*, ed. Jacqueline BROWN & William P. STONEMAN, Notre Dame (Indiana), 1997, pp. 455-504.

--------, ed., *Dating Undated Medieval Charters*, Woodbridge, Suffolk & Rochester, NY: Boydell & Brewer, and Budapest: Collegium Budapest/Institute for Advanced Study, 2000.

--------, “Identifying Irregularities and Establishing Chronology in Medieval Charters”, in K.S.B. KEATS-ROHAN, ed., *Resourcing Sources*, Occasional Publication of the Unit for Prosopographical Research, Linacre College, Oxford, vol. 7, 2002, pp. 164-78.

GURNEY, Penelope J., and Lyman W. GURNEY “Authorship Attribution: A Computer-Based Approach to a Literary Crux in Late Roman Historiography”, in *TEXT Technology*, Volume 10/1 (Winter 2001), pp. 87-105.

--------, “Enhanced content analysis of inflected languages through a system of computer-assisted lemmatization”, *Conference Abstracts of the 1994 Joint International Conference of the ALL and ACH*, Paris: The Sorbonne, pp. 93-94.

HUNYADI, Zsolt, “The Identification of a Forgery: Regularities and Irregularities in the Formulae of the Charters issued by the Székesfehérvár convent of the Knights of St. John of Jerusalem (1243-1353)”, in GERVERS, *Dating Undated Medieval Charters*, pp. 137-49.

JUHÁSZ-ORMSBY, Agnes, "Changing Legal Terminology in Dated Private Documents in England in the Twelfth and Thirteenth Centuries. A Case Study: Quitclaims", in K.S.B. KEATS-ROHAN, ed., *Resourcing Sources*, Occasional Publication of the Unit for Prosopographical Research, Linacre College, Oxford, vol. 7, 2002, pp. 195-207.

SHARPE, Richard, "Vocabulary, Word Formation and Lexicography", in: F.A.C. MANTELLO & A.G. RIGG, eds., *Medieval Latin: An Introduction and Bibliographical Guide*, Washington, D.C., 1996, pp. 93-105; "Charters, Deeds and Diplomatics", *ibid*., pp. 230-40.

SPENCER, Amanda, "Dating Charters Using Textual Evidence", in K.S.B. KEATS-ROHAN, ed., *Resourcing Sources*, Occasional Publication of the Unit for Prosopographical Research, Linacre College, Oxford, vol. 7, 2002, pp. 179-94.

XML: W3C, REC-xml-1998-02-10 "Extensible Markup Language (XML) 1.0", http://www.w3.org/TR/REC-xml

---

[1] J. Horace Round (revised and completed by W.R. Powell), "St. Thomas Becket's sisters and other studies", Essex Archaeology and History, 31 (2000), p. 154-160 (p. 156).

[2] F.M. Stenton, ed., Transcripts of Charters Relating to the Gilbertine Houses of Sixle, Ormsby, Catley, Bullington, and Alvingham, Publications of the Lincoln Record Society for 1920, 18 (Horncastle, 1922), p. x-xi, xxxiii-xxxiv.

[3] See Bibliography, 1983.