

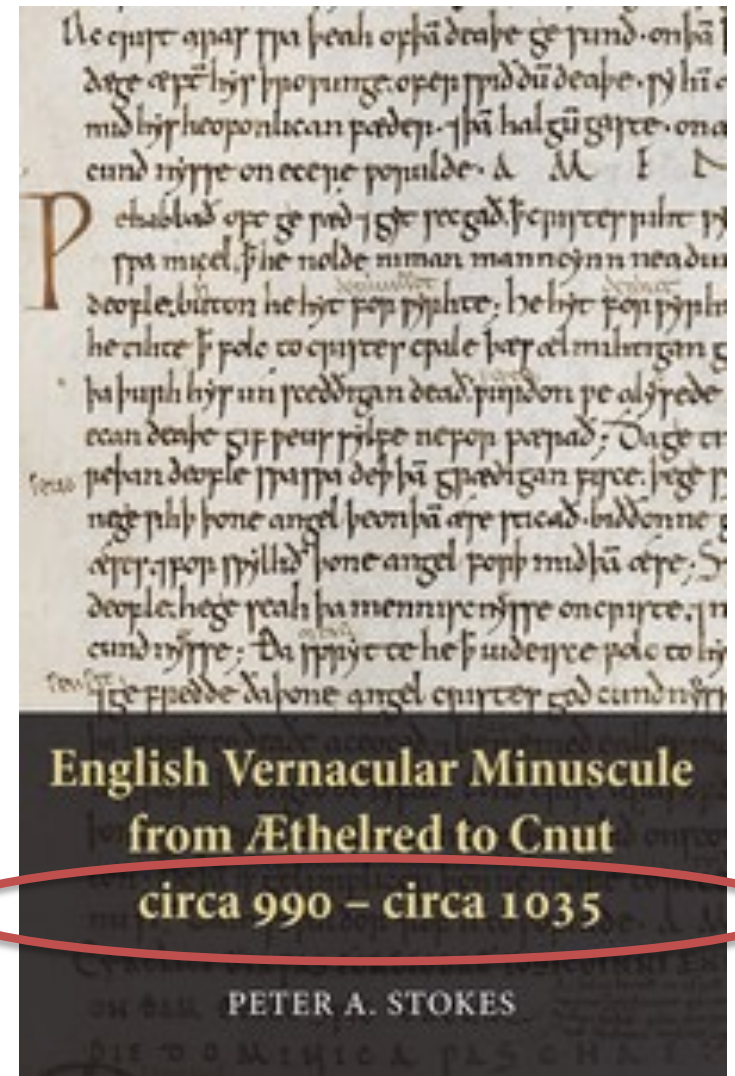
# The Problem of Digital Dating

## A Model for Uncertainty in Medieval Documents

Peter A. Stokes  
Department of Digital Humanities  
King's College, London  
[peter.stokes@kcl.ac.uk](mailto:peter.stokes@kcl.ac.uk) [@pa\\_stokes](#)  
[@digipalproject](#) [@modelsofauthority](#)

# The Challenge from a Palaeographer

My challenge: Go to an online catalogue of manuscripts and, in a single query, find me all manuscripts containing writing dated circa 990 – circa 1035.



# Uncertainty and Expectations in Dating

---

- Many manuscripts and documents can't be dated exactly
- Dating normally imprecise: 'early twelfth century', etc.
  - Scholarly convention: *saec. xii in.*, *saec.*, *xii<sup>2</sup>*, etc.
- This has worked (fairly) well to date, but doesn't translate well digitally:
  - Is 'early twelfth century' before 'first quarter of twelfth century'?
  - Does 'first half of twelfth century' include 1150? 1151? 1000? 1099?
  - Does '690 (or 960?)' come before or after 'circa 900'
  - Should a search for 'mid-ninth century' include 'circa 835'?
- Question largely not addressed in existing sites (I think!):  
*How to design interface to search, sort, arrange by date?*

# A Straw Man? The Survey

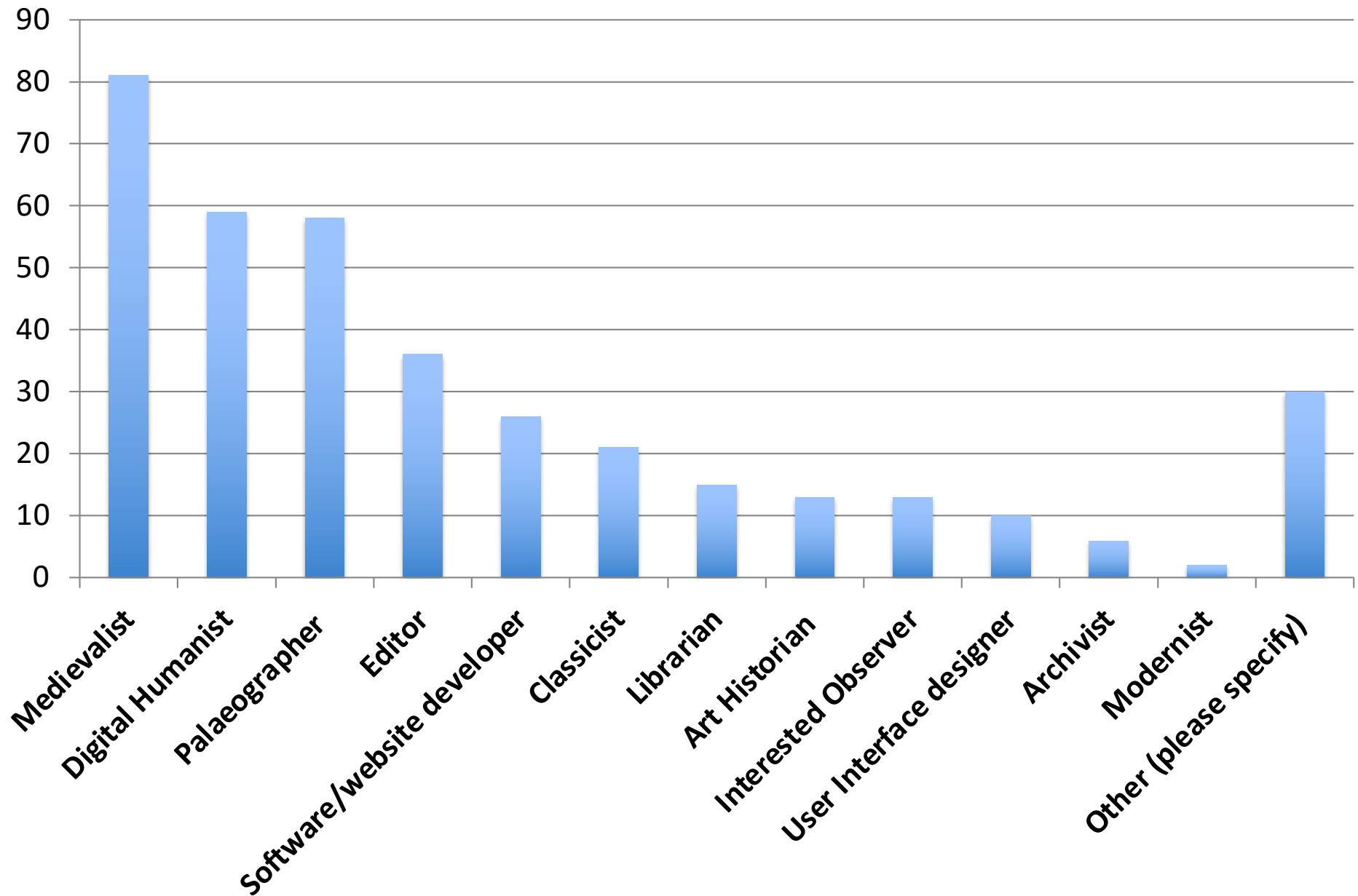
---

Numerous people have suggested this isn't a real problem:

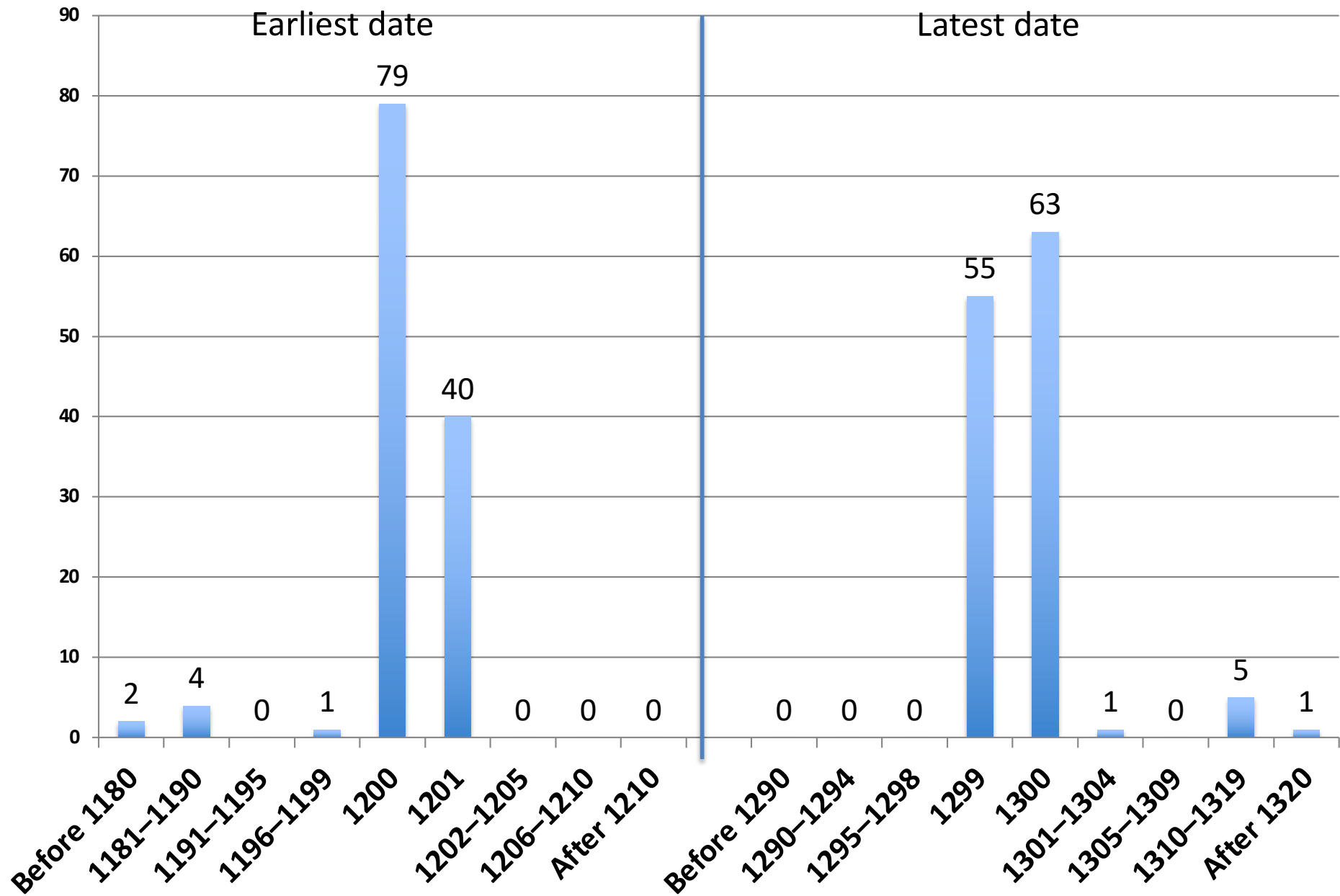
- We know when 12<sup>th</sup> century began: in 1101 (see Wikipedia)
  - Or in 1100 (by convention)
- We know the 12<sup>th</sup> century began on 1 January
  - Or on Christmas Day, or on Lady Day, or on...
- Scholarly convention is clear that 'early' = first decade
  - Or first quarter-century, or first third, or ...
- It doesn't matter: we just need to agree a convention
- It doesn't matter: the question is ill-posed with no answer

-> Survey by Matthew Driscoll and myself to judge convention

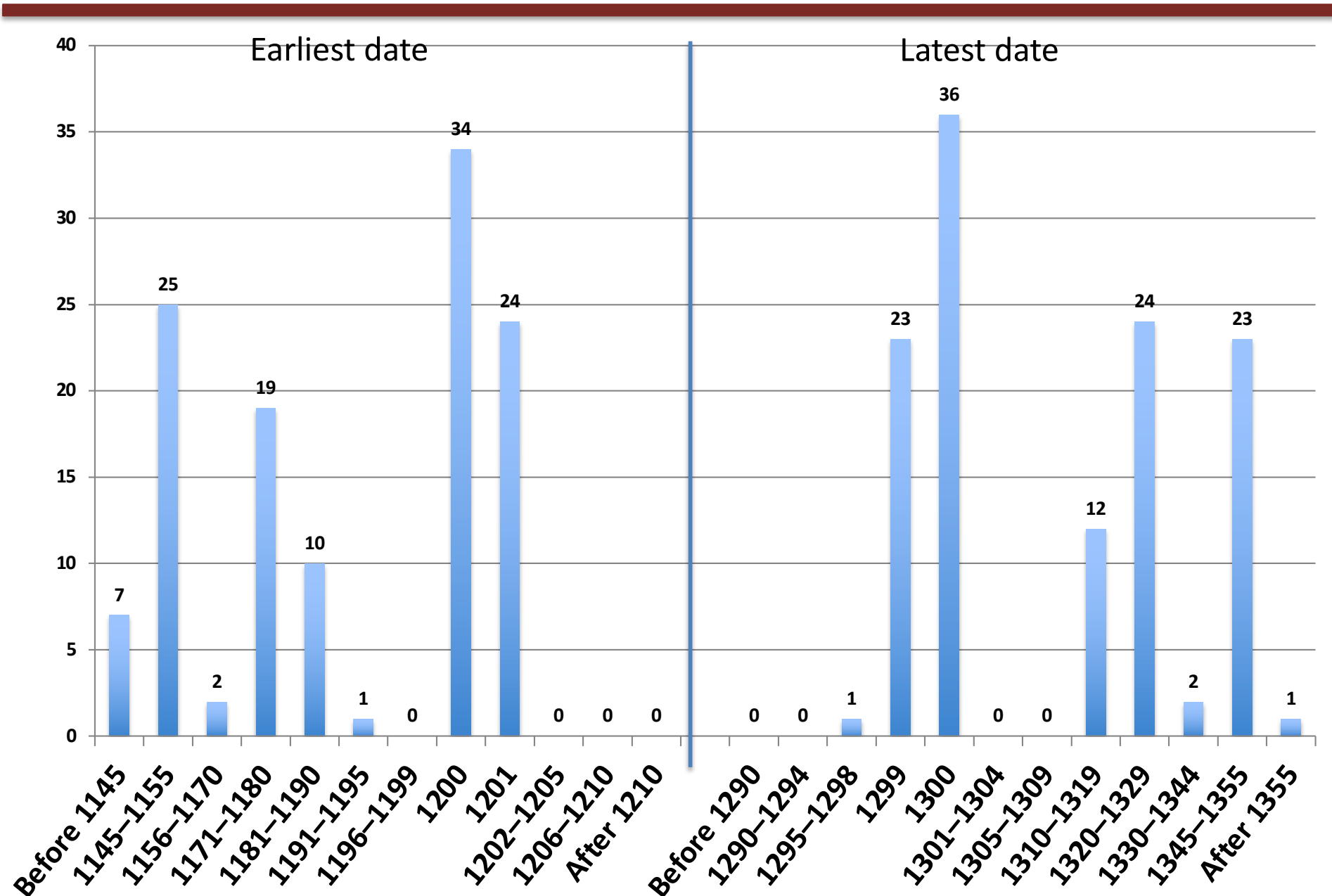
# Which of the following best describe you? (Check all that apply)



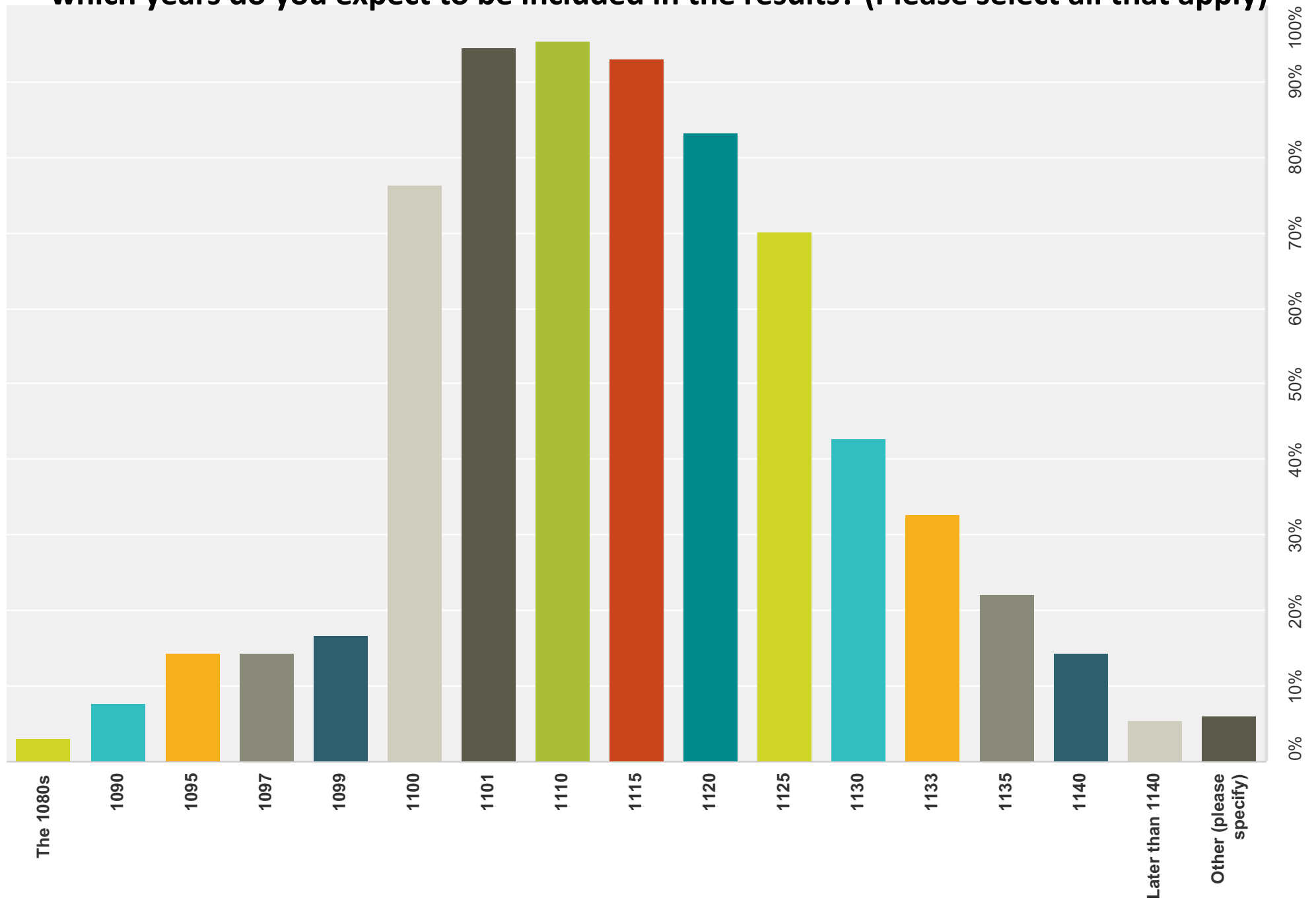
When searching a database or on-line catalogue for material dated "thirteenth century", what range of dates would you expect to be included in the results?



When searching a database or on-line catalogue for material dated "thirteenth century (?)", what range of dates would you expect to be included in the results?

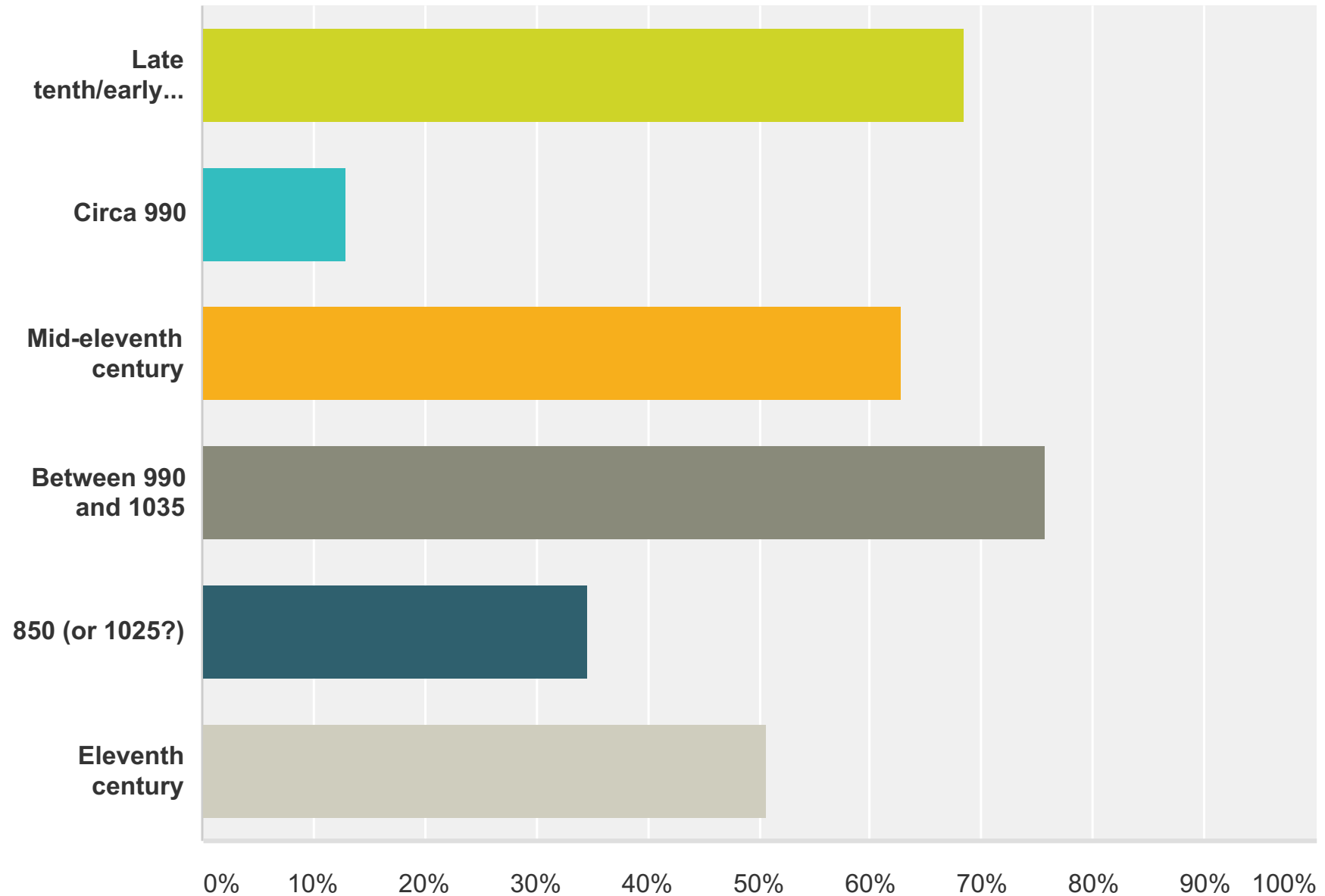


**When searching a database or on-line catalogue for "early twelfth century" (saec. xii in.), which years do you expect to be included in the results? (Please select all that apply)**

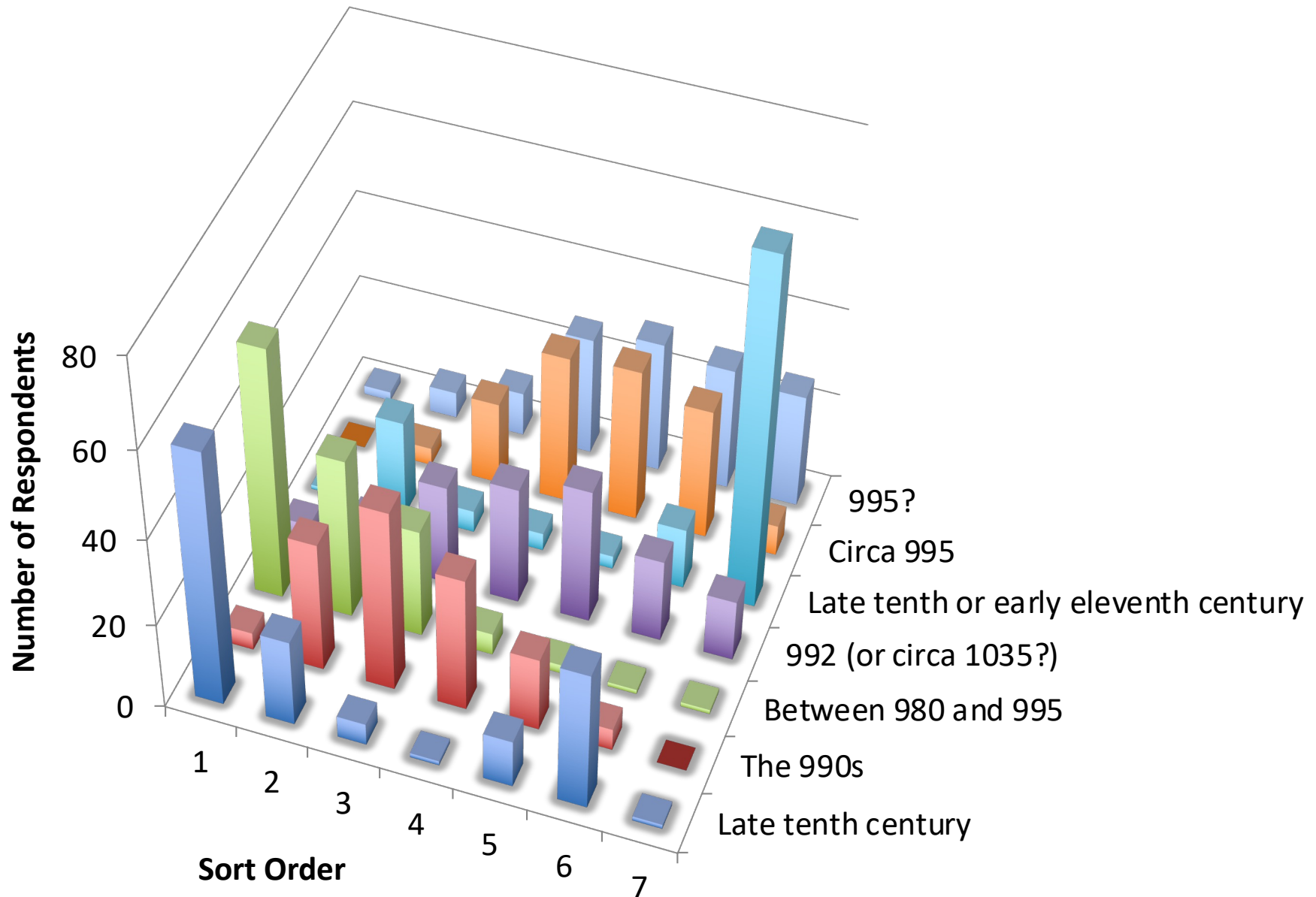




**When searching for material from the "first half of the eleventh century", which date-range(s) would you expect to have returned in your results? (Check all that apply.)**



Please sort the following into chronological order  
(1 for earliest through to 7 for latest)

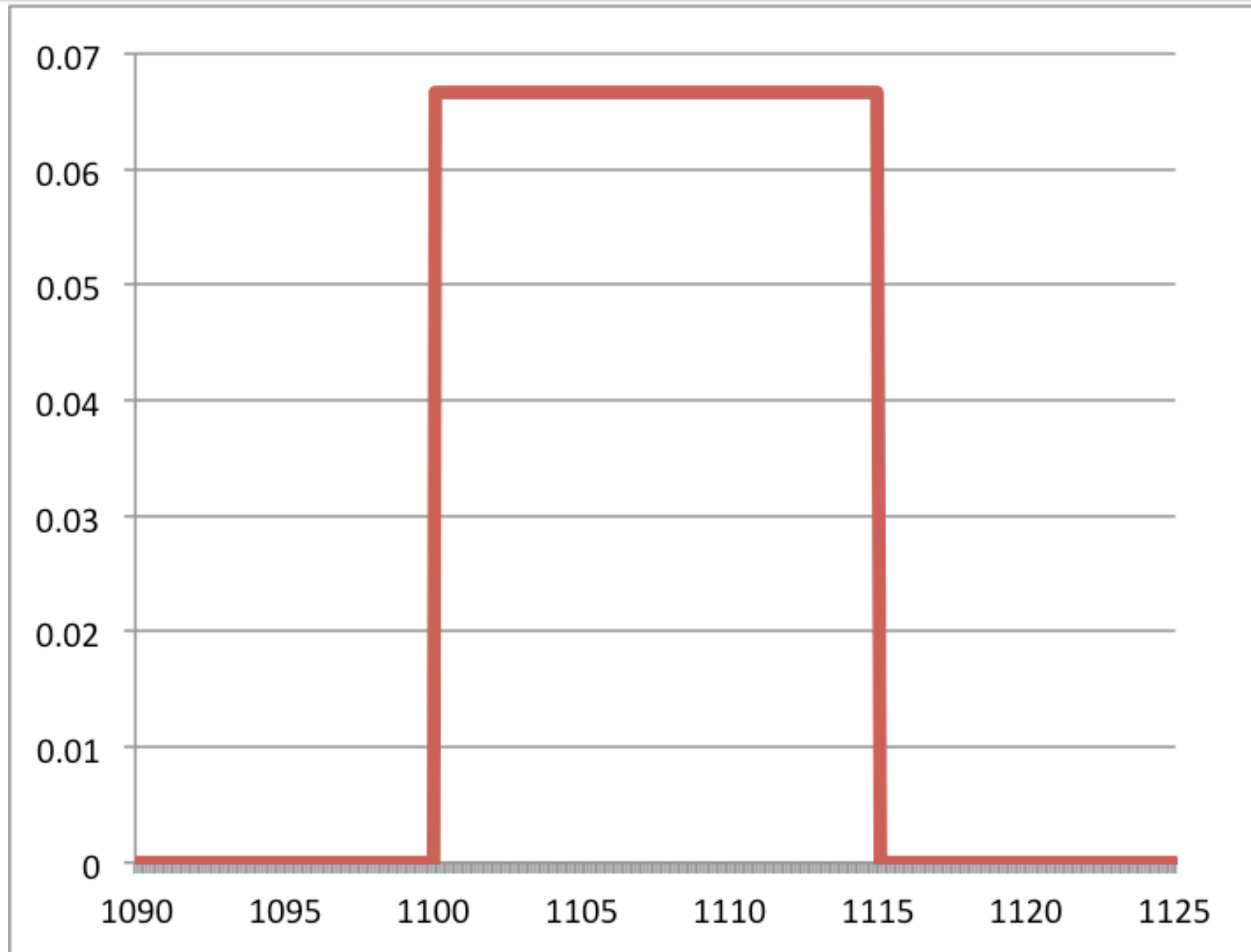


# Solution? We need to start again...

---

- In this context, ‘early twelfth century’ does not mean *and has never meant* ‘not before 1 January 1100 and not after 31 December 1110’ (or 1115, or 1120, or...)
- Something much closer to ‘*probably not much* before 1100, and *probably not much* after 1110 [or whatever], but *probably* somewhere in between’.
- We’re normally dealing with probabilities anyway: if we knew the date then we (probably) wouldn’t say ‘early twelfth century’.
- *If these dates are probabilities anyway, then why aren’t we treating them accordingly?*
  - Cf. ‘temporal geometry’ of Grossner and Meeks

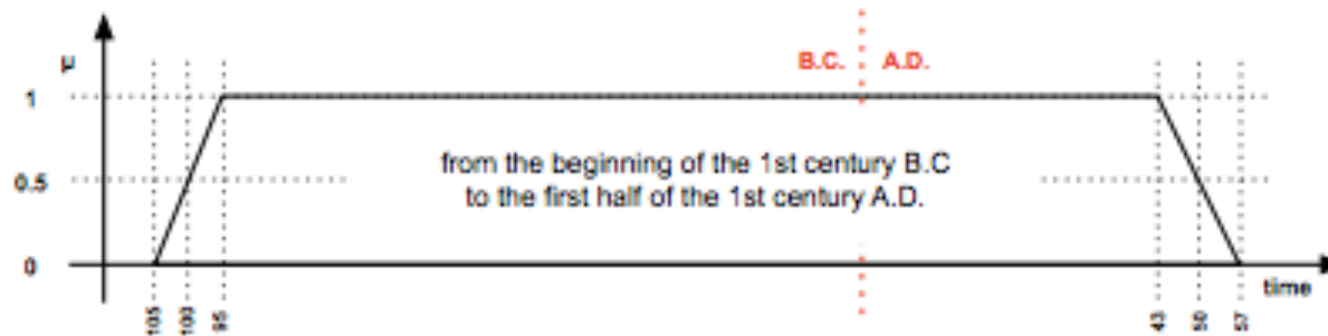
# 'During early 12<sup>th</sup> century' to a computer?



# Time period to fuzzy logicians (Kauppinen et al.)

---

$$\begin{aligned} T_{fuzzybegin} &= 105 \text{ B.C.} \\ T_{begin} &= 95 \text{ B.C.} \\ T_{end} &= 43 \text{ A.D.} \\ T_{fuzzyend} &= 57 \text{ A.D.} \end{aligned}$$



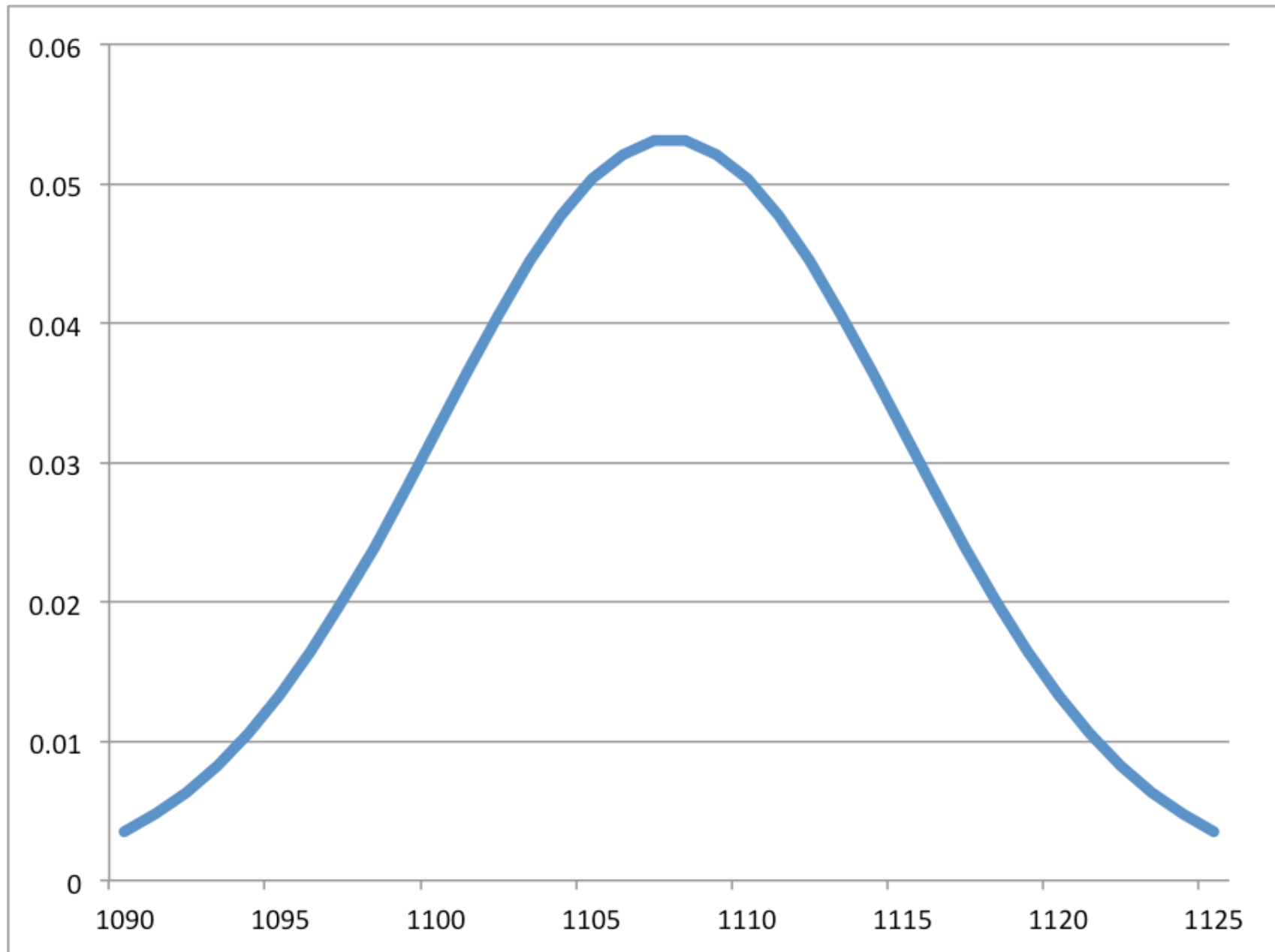
From T. Kauppinen et al. (2010), 'Determining relevance of imprecise temporal intervals ...'  
*Int. Jnl of Human-Computer Studies* 68 (2010) 549–560

Cf. TEI, Simile, TopoTime, ...

---

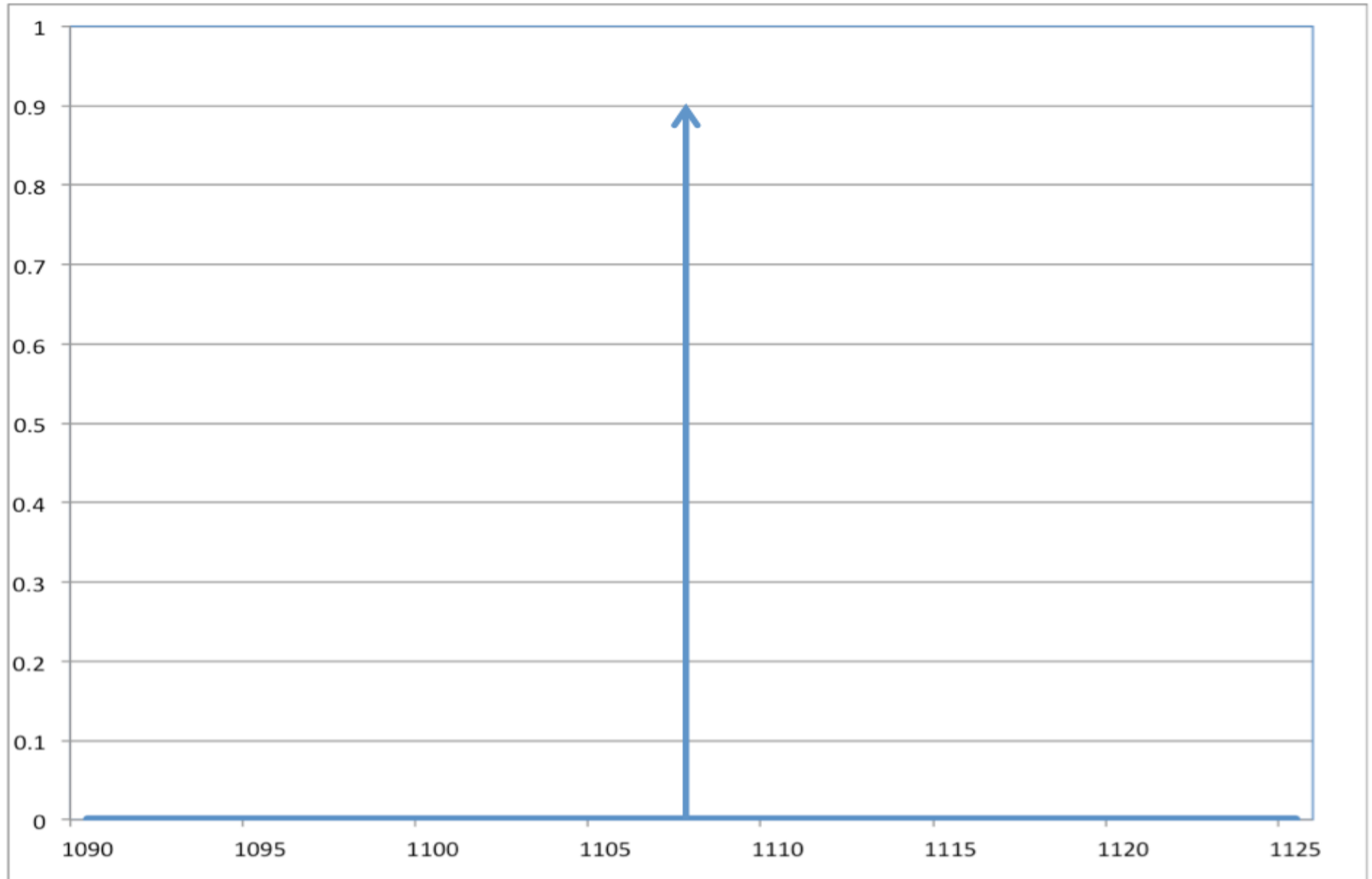
# 'During early 12<sup>th</sup> century' to a medievalist?

---



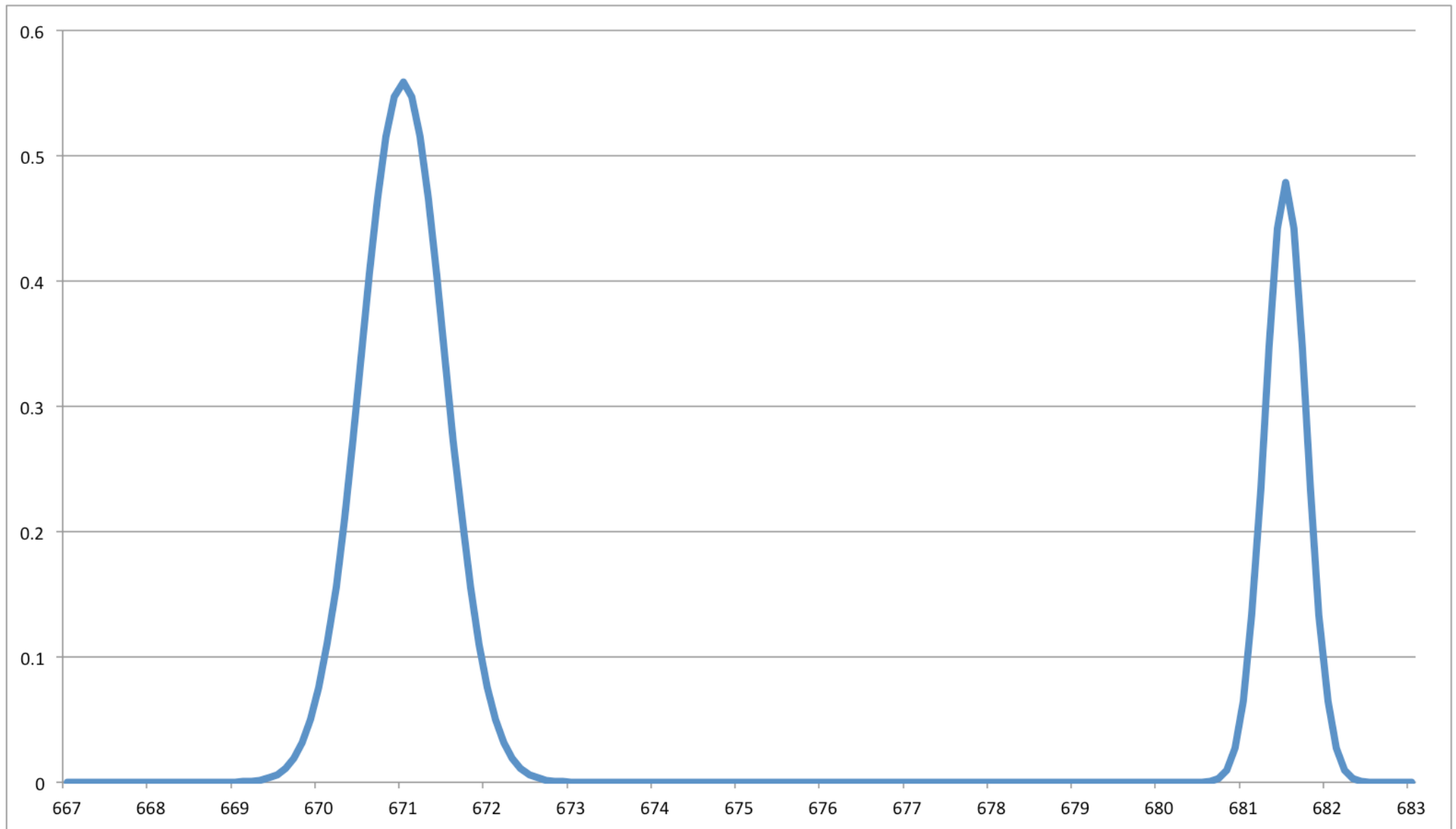
# Dated 1107

---



# 670x672 (or 681?)

---



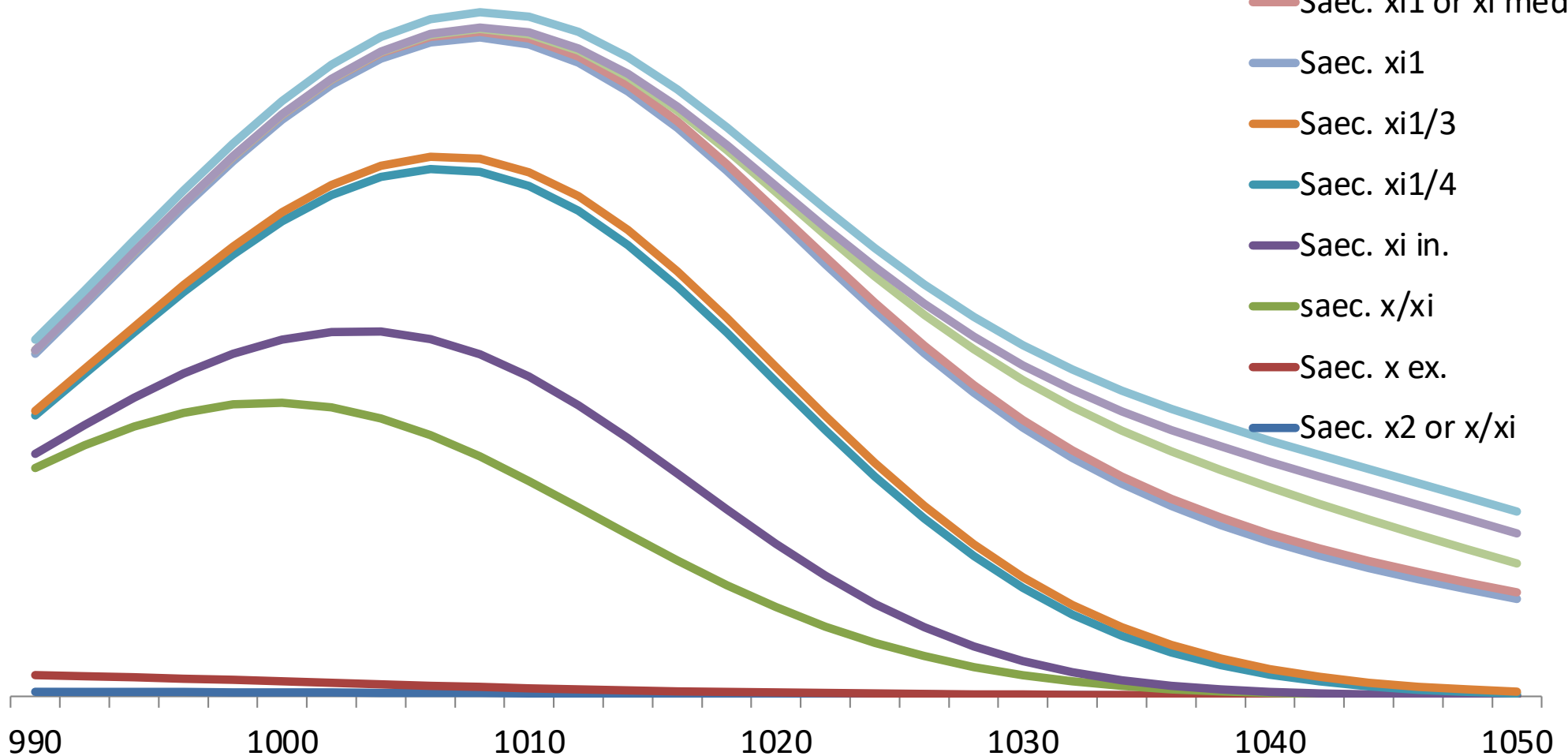


# Distribution of Tall-e æ in Early 11<sup>th</sup> Cent.

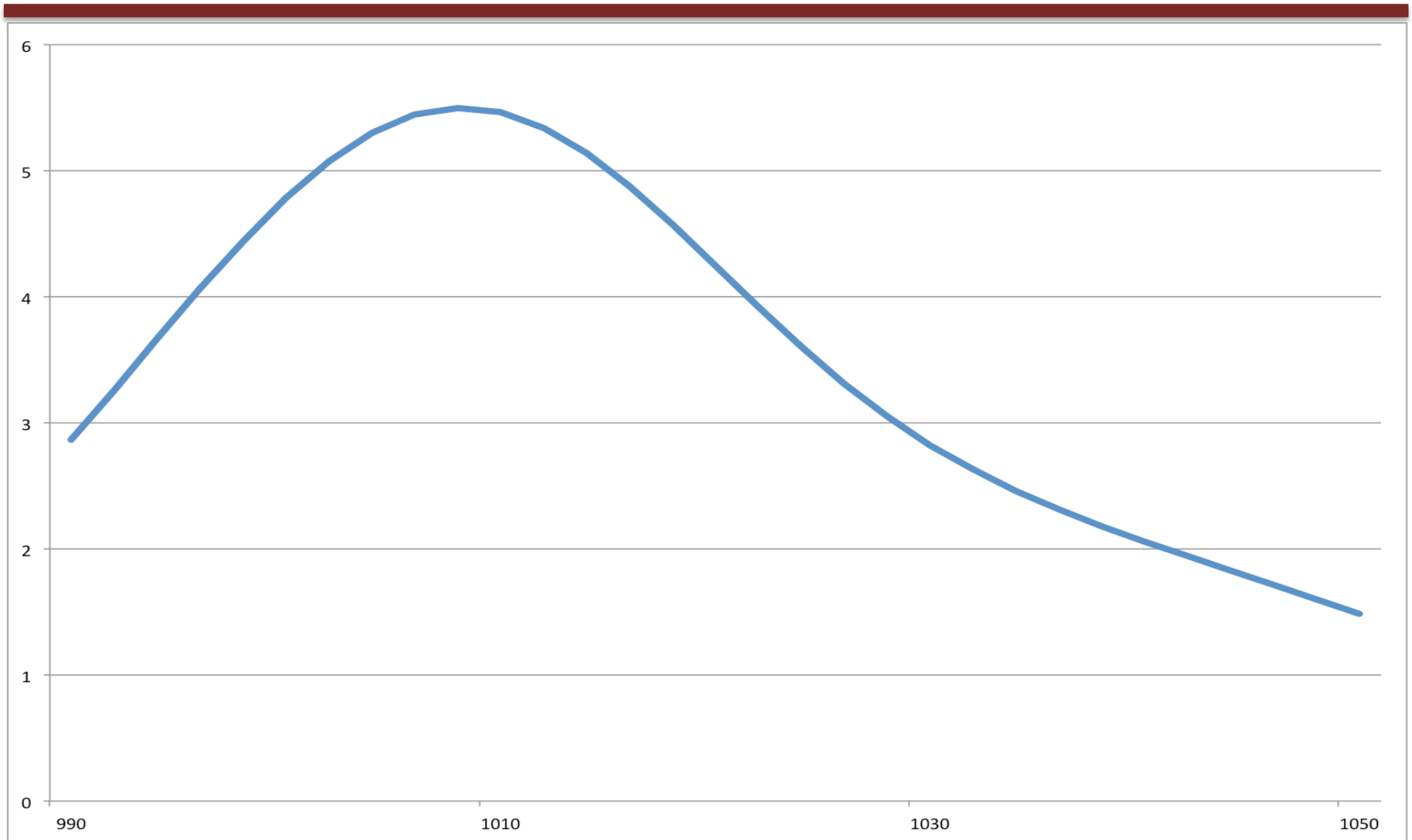
Weight each curve by number of occurrences, then sum to produce aggregate 'distribution'



- Saec. xi
- Saec. xi med.
- Saec. xi2/4
- Saec. xi1 or xi med.
- Saec. xi1
- Saec. xi1/3
- Saec. xi1/4
- Saec. xi in.
- saec. x/xi
- Saec. x ex.
- Saec. x2 or x/xi



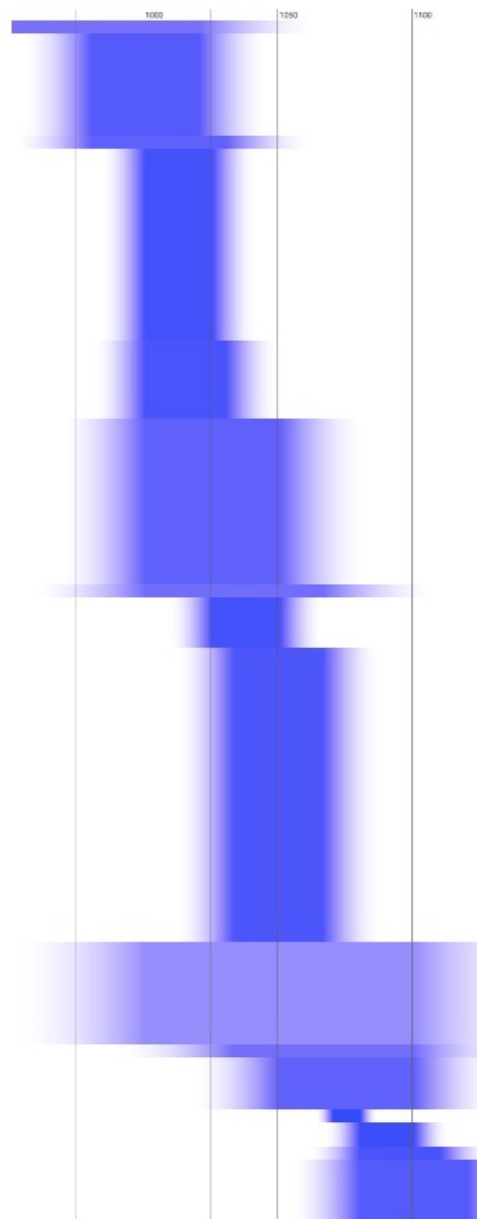
# Distribution of Tall-e æ in Early 11<sup>th</sup> Cent.



Works well enough for a single aggregate,  
but what about comparison across categories?

# Dist. of Scribal Hands from Canterbury CC

95 matching results



- BL Royal 12.C.xviii [Twenty-eight Glosses] (Sae. x2 or x1x)
- OCDD 326 [Adhelm: Poem (pp. 5-6)] (Sae. x1x)
- OCDD 328 [Main Glossator (approx. 86 glosses)] (Sae. x1x)
- BL Cotton Otho 2 [Additional Lemmata and Glosses] (Sae. x1x)
- BL Royal 6.A.v [Glossator 1] (Sae. x. ex. or x1x)
- Bookeian Bookley 577 (27645) [Glossator 1] (v. 35v, 36r, 65v, 71r, and 73v) (Sae. x1x)
- BrF latn 6421A [Latin Gloss] (Sae. x. ex. or x1x)
- OCDD 173, fol. 1-56 [Addition (Annal for 11AD)] (Sae. x1x)
- OCDD 173, fol. 1-56 [Addition to annals up to 616] (Sae. x1x)
- BL Royal 6.A.v [Hemle's Two Glosses (32v-48r)] (Sae. x1x or x1x)
- BL Royal 1.D.1x [Confirmation of Privileges (44v)] (Sae. x11x)
- BL Slove Ch. 37 [Hand 1 (lines 1-6)] (Sae. x11x)
- BL Slove Ch. 37 [Hand 2 (lines 7-26)] (Sae. x11x)
- BL Slove Ch. 38 [Main Hand] (Sae. x11x)
- Canterbury Chart Ant. x108 [Main Hand] (Sae. x11x)
- BL Slove Ch. 2 [Main Hand (Latin)] (Sae. x11x)
- TCC D.10.4 (215) [Main Hand (Latin)] (Sae. x11x)
- BL Additonal 34900 [Main Hand (Latin)] (Sae. x11x)
- BL Anundel 155, fol. 1-135 and 171-91 [Latin Hand] (Sae. x11x)
- BL Cotton Vespasian C.ii, fol. 11-85 [Main Hand] (Sae. x11x)
- BL Royal 1.D.1x [Notice of Canterbury (32v)] (Sae. x11x)
- BL Royal 1.D.1x [Latin Hand 1] (Sae. x11x)
- BL Royal 1.D.1x [Latin Hand 2] (Sae. x11x)
- Rowan, DM 274 (74) [Names of months (fol. 6-11), Rubrics (fol. 207-12), Rubrics (fol. 207-12)] (Sae. x11x)
- BL Slove Ch. 38 [Main Hand (Latin)] (Sae. x11x)
- OCDD 146 [Adulfur's (pp. 302-308)] (Sae. x11x)
- BL Additonal 37517 [Continuous Gloss] (Sae. x11x)
- BL Royal 5.E.ii [Six Glosses A (34r glosses, mostly on 271-115v)] (Sae. x11x)
- BL Royal 5.E.ii [Six Glosses B (awynia, genit. on 115v, welf, ream, win, wine, 117v)] (Sae. x11x)
- Bookeian Bookley 163 (2014), fol. 1-227, 252 and 251 [Main Latin Hand] (Sae. x11x)
- OCDD 146 [Main Latin Hand (pp. 61-316)] (Sae. x11x)
- Lambeth 1370 [Boundaries (114v, 5, 156r)] (Sae. x11)
- Lambeth 1370 [Writ of Out (114v, 5, 156r)] (Sae. x11)
- Lambeth 1370 [Record of Consecration (89v, 8, 138r)] (Sae. x11)
- BL Anundel 155, fol. 1-135 and 171-91 [Continuous Glossator 1 (171-182)] (Sae. x11)
- BL Anundel 155, fol. 1-135 and 171-91 [Continuous Glossator 2 (182-182)] (Sae. x11)
- BL Cotton Tiberius A.ii, fol. 2-173 [DE Hand 1 (glosses on 2v-27v and 80v-84v; additions on 37v-42v and 85v-89v)] (Sae. x11)
- BL Cotton Tiberius A.ii, fol. 2-173 [DE Hand 2 (glosses on 27v-37v and 711v-88r; text on 46v-56v and 99v-107v)] (Sae. x11)
- BL Royal 2.B.v [Addition Hand 4 (167-88)] (Sae. x11)
- Boulogne 9M 189 [Glossator 1 (fol. 4-11, 90, 102)] (Sae. x11)
- Boulogne 9M 189 [Glossator 2 (fol. 74-82)] (Sae. x11)
- Boulogne 9M 189 [Glossator 3 (fol. 100-104, 111-122, 139-43, 180)] (Sae. x11)
- Boulogne 9M 189 [Glossator 4 (throughout)] (Sae. x11)
- Geneve 175 [One gloss (bybonat, 11v10)] (Sae. x11)
- BL Cotton Julius A.ii [Continuous Glossator] (Sae. x11 or x1 med)
- BL Cotton Tiberius B.ii, fol. 87 [Writ of Out (5, 887 and 888)] (Sae. x21x)
- BL Slove Ch. 38 [Continuation] (Sae. x21x)
- BL Slove Ch. 42 [Main Hand] (Sae. x21x)
- BL Cotton Tiberius B.ii, fol. 2-73 and 77-88 [Main Hand (incl. scribbles)] (Sae. x21x)
- BL Cotton Augustus 8.70 [Main Hand] (Sae. x1 med)
- BL Cotton Augustus 8.70 [Main Hand] (Sae. x1 med)
- BL Cotton Claudius A.ii, fol. 2-7 and 37 [Record A (div. from 71c wile)] (Sae. x1 med)
- BL Cotton Claudius A.ii, fol. 2-7 and 37 [Record B (div. from 71c wile)] (Sae. x1 med)
- BL Cotton Claudius A.ii, fol. 2-7 and 37 [Record C (div. from 71c wile)] (Sae. x1 med)
- BL Cotton Claudius A.ii, fol. 2-7 and 37 [Record D (div. from 71c wile)] (Sae. x1 med)
- BL Cotton Tiberius A.ii, fol. 2-173 [DE Hand 3 (fol. 43-48)] (Sae. x1 med)
- BL Cotton Tiberius A.ii, fol. 174-7 [DE Hand] (Sae. x1 med)
- BL Cotton Vespasian D.ii [Hand 1, incl. cont. gloss (80v-157)] (Sae. x1 med)
- BL Cotton Vespasian D.ii [Hand 2, incl. cont. gloss (80v-157)] (Sae. x1 med)
- BL Royal 2.B.v [Addition Hand 5 (8)] (Sae. x1 med)
- BL Royal 2.B.v [Addition Hand 6 (8)] (Sae. x1 med)
- BL Royal 5.E.ii [Sixty Glosses (mostly 11v-115v)] (Sae. x1 med)
- BL Royal 7.C.1v [Main Glossator] (Sae. x1 med)
- BL Royal 7.C.1v [Earlier Glossator (fol. 24, 31, 25)] (Sae. x1 med)
- Bookeian Ashmole 328 (882 and 740) [Main Hand] (Sae. x1 med)
- Worcester Q.2 [Reading (76v)] (Sae. x1 med)
- BL Cotton Tiberius A.ii, fol. 2-173 [Division to Prayers, Chant] (Sae. x1 med)
- BL Cotton Augustus 8.57 [Main Hand] (Sae. x1)
- Canterbury Chart Ant. C.1279, Red Book, no. 7 [Main Hand] (Sae. x1)
- BL Cotton Vespasian A.ii [Added Glosses] (Sae. x1)
- BL Royal 2.B.v [Iterations (190v-96v)] (Sae. x1)
- BL Royal 2.B.v [Scribbles (fol. 188v)] (Sae. x1)
- BL Royal 2.B.v [Scribbles (fol. 188v and 189v)] (Sae. x1)
- Boulogne 9M 189 [Scratched Gloss] (Sae. x1)
- TCC O.2.31 (115) [Scratched Glosses] (Sae. x1)
- BL Royal 2.B.v [Office of the Virgin (11-6)] (Sae. x1 med. or x12)
- BL Cotton Claudius A.ii, fol. 2-7 and 37 [Record E: Writ of Edward the Confessor (3v)] (Sae. x12)
- BL Cotton Ch. x1.31 [Main Hand] (Sae. x12)
- OCDD 173, fol. 1-56 [Chronicle Hand 6 (a, 107r)] (Sae. x12)
- TCC B.15.34 (268) [Connections and Alterations] (Sae. x12)
- BL Cotton Caligula A.ii, fol. 125-43 [Main Hand (Annals, etc.)] (Ca 1075)
- BL Slove Ch. 42 [Main Hand] (Sae. x1)
- BL Cotton Tiberius A.ii, fol. 2-173 [Annalator] (Sae. x1 ex.)
- Rylands 109 [Latin Glosses] (Sae. x1 ex. or x1v)
- CUL 8.2.1 (1734) [Latin Gloss] (Sae. x1v)
- OCDD 173, fol. 1-56 [List of Popes and Archbishops of Ca (54v, Ker article 4)] (Sae. x1v)
- BL Cotton Caligula A.ii, fol. 125-43 [Annal 1 (a, 106v-77r)] (Sae. x1v)
- BL Cotton Domitian v.ii, fol. 30-70 [Main Hand] (Sae. x1v)
- OCDD 173, fol. 1-56 [Addition (a, 106r)] (Sae. x1v)



# Search Results

---

- Don't define by strict numerical range, *or* by text
- Define by significance, according to overlap within range
  - 'Query timespan and return a list of the periods in the collection that overlap with it, ordered by the area of intersection' (Grossman & Meeks)
  - Mathematically: integrate pdf over date-range to get significance
  - Works for true pdf, not for fuzzy logic as per Grossman & Meeks
  - Can calculate simply with XSLT using lookup tables
- I'm not convinced the ordering is helpful
  - Very difficult for users to understand
  - Seems better to order by mid-point, or perhaps peak of pdf

# What do these Numbers Represent?

---

These do not represent 'real' likelihoods and must be used with great caution!

- How can I measure the 'real' probability of my judgment?
- How can I measure the probability of another's judgment?
- How can I measure the combination of others' judgments?
  - Remember, these dates can come from different sources
- I would not use these to date unknown hands, manuscripts
- My goal is to provide more intuitive and useful searching

# Thanks to

---

Matthew Driscoll for co-authoring the survey

The DigiPal & MoA Teams (see website for details)

The European Research Council and the  
Arts and Humanities Research Council (UK)  
for funding this research

<http://digipal.eu/>

@digipalproject @modelsofauthority

*The research leading to these results has received funding from the European Union  
Seventh Framework Programme (FP7) under grant agreement n° 263751.*

*This work is now supported by the Arts and Humanities Research Council (UK),  
grant nos. AH/L008041/1 and AH/L013975/1.*