Hello!

This is a corpus of plural forms (and morphological collectives) collected from Old Welsh and a sample of Middle Welsh texts. The latter are divided into three groups: poetry, native prose and translated prose (from Latin). The first version of the corpus was collected for my 2010 MPhil thesis and contained all relevant Old Welsh forms (the Old Welsh corpus as a whole is small) and a sample of Middle Welsh poetry. To this, the Middle Welsh prose corpus was added and this formed the basis for my PhD thesis (2015) and subsequent monograph (2019):

Nurmio, S. 2010. Astudiaeth o eiriau Cymraeg a chanddynt fwy nag un ffurf luosog mewn testunau Hen Gymraeg a Chymraeg Canol Cynnar. Aberystwyth University MPhil dissertation. http://hdl.handle.net/2160/5086

Nurmio, S. 2015. Studies in Grammatical Number in Old and Middle Welsh. University of Cambridge PhD dissertation.

Nurmio, S. 2019. *Grammatical Number in Welsh: Diachrony and Typology*. Publications of the Philological Society 51. Malden, MA, Oxford and Chichester: Wiley & Blackwell. https://onlinelibrary.wiley.com/toc/1467968x/2019/117/S1

Here are some details on the different parts of the corpus:

1. Poetry and Old Welsh: this is the sample originally collected for my MPhil thesis and it is therefore a little different in format from the other two databases. Especially for Old Welsh, some lexemes are very unclear, but if they seem likely to be plural forms, they are included with '?' in the 'Singular' column and with notes if there are any available. The practice with the poetry from *Cyfred Beirdd y Tywysogion* (*CBT*) is that also used in the MPhil: if there are many references, only the first is listed in the 'Text' column while all the poets using the form are listed in the 'Beirdd y Tywysogion' column, allowing the reader to check all references if needed

2. Native prose: no page references are given as all these texts are available in the *Rhyddiaith* corpus (www.rhyddiaithganoloesol.caerdydd.ac.uk) by Luft et al. (2013) and can be searched for quickly. An exception is *Breudwyt Ronabwy* where page numbers in the edition by Richards are given.

3. Translated prose: references are to folia in Parry's edition (for Cotton Cleopatra Bv), and to pages in Isaac et al. (2013), *Rhyddiaith Gymraeg o Lawysgrifau'r 13eg Ganrif Fersiwn 2.0* (www.academia.edu/610554/Rhyddiaith_Gymraeg_o_Lawysgrifaur_13eg_Ganrif_Fersiwn_2.0) (for Llanstephan 1 and Peniarth 44). For Peniarth 44, some references are to Roberts (1969) in which case the format is [page].[line], e.g. 11.18 means 'page 11, line 18'. For *Brut Dingestow* there are also two types of references: those to Isaac et al. (2013) are to a page number only, while those to Lewis, *Brut Dingestow* are in the format [page].[line]

Here is a list of the texts used in the corpus, with abbreviations:

OLD WELSH

Ang	Angers MS 477 (see Falileyev 2000, *Etymological Glossary of Old Welsh*: 14)
Chad 2	the 'Surexit' Memorandum in the Book of St. Chad (see Falileyev 2000: 14)
J	glosses in Cambridge MS Ff 4. 42 (see Falileyev 2000: 15)
Juv	9 nine englynion in Cambridge MS Ff 4. 42 (see Falileyev 2000: 16)

LL	Evans & Rhŷs (eds.). 1893. *The text of the Book of Llan Dav*
MC	glosses on Martianus Capella, De Nuptiis Philologiae et Mercurii, Corpus Christi College, Cambridge MS 153 (see Falileyev 2000: 16)
MP	glosses on De mensuris et ponderibus (Liber Commonei, Oxoniensis Prior), Bodleian MS. Auct. F. 4. 32 (see Falileyev 2000: 16)
Ovid	glosses on Ovid, Ars Amatoria (Oxoniensis Prior, Bodleian MS. Auct. F. 4. 32, fol. 37a – 42a) (see Falileyev 2000: 16)
Ox2	glosses on "Oxoniensis posterior" (Bodleian MS 572) (see Falileyev 2000: 17)
Pad	glosses on Corpus Christi MS 199 fo. 11a (see Falileyev 2000: 17)
Prisc	glosses on Priscian's Grammar (BN lat. 10290) (see Falileyev 2000: 17)

MIDDLE WELSH

Poetry

CA	Williams (ed.). 1961. *Canu Aneirin*.
CBT	Gruffydd (series ed.) et al. *Cyfres Beirdd y Tywysogion*, 7 vols (referred to as CBT 1, CBT 2, etc.).
LlDC	Jarman (ed.). 1982. *Llyfr Du Caerfyrddin*.

Native Prose

4 Branches	Luft et al. (2013). The four branches of the Mabinogi.
Gereint	Thomson (ed.). 1997. *Gereint fab Erbin*. Also in Luft et al. (2013).
CO	Bromwich & Evans (ed.). 1997. *Culhwch ac Olwen*. Also in Luft et al. (2013).
Ll&Ll	Roberts (ed.). 1975. *Cyfranc Lludd a Llefelys*. Also in Luft et al. (2013).
Maxen	Roberts (ed.). 2005. *Breudwyt Maxen Wledic*. Also in Luft et al. (2013).
Owein	Thomson (ed.). 1968. *Chwedyl Iarlles y Ffynnawn*. Also in Luft et al. (2013).
Peredur	Goetinck (ed.). 1976. *Historia Peredur vab Efrawc*. Also in Luft et al. (2013).
Rhonabwy	Richards (ed.). 1948. *Breudwyt Ronabwy*. Also in Luft et al. (2013).

Translated Prose

BD	NLW 5288 (*Brut Dingestow*), edited both in Isaac et al. (2013), and Lewis, 1942. *Brut Dingestow*.
BB1	British Library MS Cotton Cleopatra Bv, ed. Parry, 1937. *Brut y Brenhinedd: Cotton Cleopatra version*.
Llanst 1	NLW Llanstephan 1, ed. Isaac et al. (2013)
Pen 44	NLW Peniarth 44, ed. Roberts, B. F. 1969. Astudiaeth destunol o'r tri chyfieithiad Cymraeg cynharaf o *Historia Regum Britanniae* Sieffre o Fynwy: ynghyd ag "argraffiad" beirniadol o

destun Peniarth 44. http://cadair.aber.ac.uk/dspace/handle/2160/5869; also in Isaac et al. (2013)

Latin references are to Reeve (ed.) & Wright (trans.) 2007. *Geoffrey of Monmouth: The History of the Kings of Britain; an Edition and Translation of* De Gestis Britonum (Historia Regum Britanniae). Woodbridge: Boydell Press.

Here is a list of other abbreviations that appear in the corpus:

adj	adjective
attr	attributive adjective
BO	'*Brutiau* only'; form attested only in the *Brut y Brenhinedd* texts (based on attestations given in *Geiriadur Prifysgol Cymru*)
BR	'*Brutiau* + rare'; form attested in the *Brut y Brenhinedd* texts and rare elsewhere (attestations given in the next column)
cmpd	compound noun
coll	morphological collective (noun)
coll & pl	plural form based on a morphological collective, e.g. *coed-ydd* 'forests'
cyh. hir	*cyhydedd hir* (rhyme)
edig	derivative in *-edig*
gloss	(for Old Welsh) the noun appears as an interlinear gloss on a Latin text
i-aff	(plural formed with) i-affection (aka vowel alternation), e.g. *march* 'horse', pl. *meirch*
/j/	for this noun, plurals with and without a glide or 'yod' are attested in the sample, e.g. *meibion* and *meibon* 'sons'
num	numerative (form used after numerals, may be identical with the plural of that noun)
pl & pl	double plural form (or appearing so synchronically), e.g. *gwreich-ion-au* 'sparks'
pl tantum	plurale tantum; a noun which does not have a singular form, e.g. *rhieni* 'parents'
pred	predicative adjective
Proph	*Prophetiau Merlini*, a section found in the *BD* and *BB₁* versions of *Brut y Brenhinedd*
subst	adjective used substantivally
-wyr	plural form with the ending *-wyr* (from *gwŷr* 'men')

If the lemma in the Singular column is capitalised, this means that the noun has different plural forms in that corpus, e.g. *sant* 'saint', plurals *seint, sein(i)au* in the Poetry corpus.