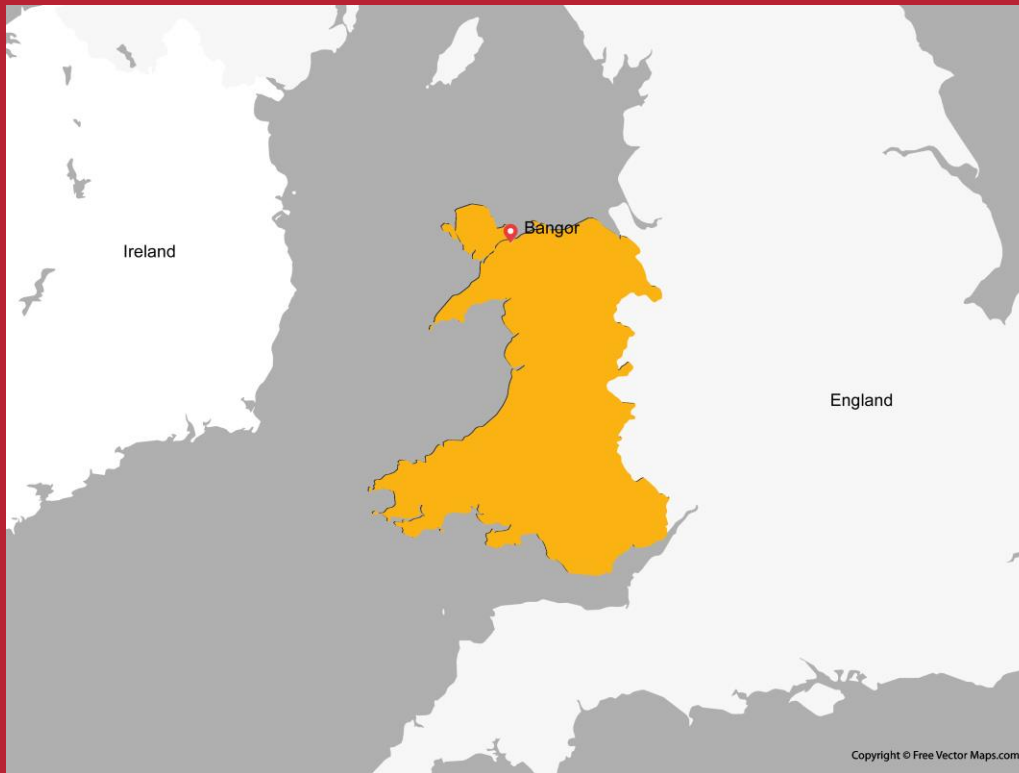


Delyth Prys, Mared Roberts, Gruffudd Prys
Language Technologies Unit, Canolfan Bedwyr, Bangor University

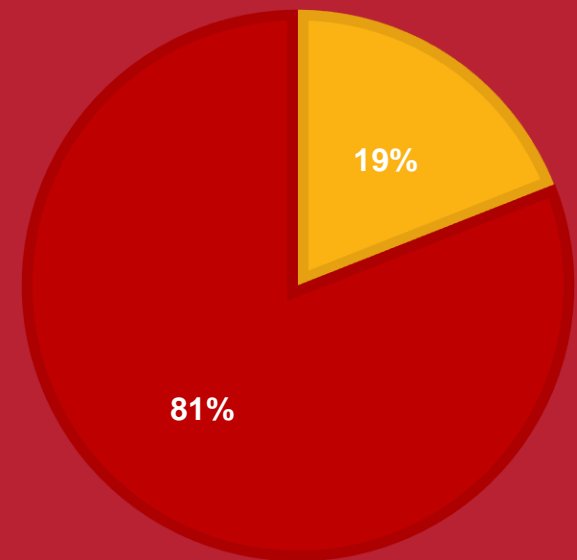
2nd Workshop on Collaboration and Computing for Under-Resourced Languages, 23 May 2016, LREC Portorož

Context: Wales and Welsh

Wales: Population 2,955,841*



19% speak Welsh*
(562,016)



***2011 Census**

See: <https://statswales.wales.gov.uk/Catalogue/Welsh-Language/WelshSpeakers-by-LocalAuthority-Gender-DetailedAgeGroups-2011Census>

Context: Book publishing in Wales

- Subsidies and grants are used to make publishing books commercially viable
- The Welsh Books Council is a national body, funded by the Welsh Government
- 367 Welsh language books were published in 2015 (excluding textbooks)
- Few grants to cover reprinting costs

Context: Coleg Cymraeg Cenedlaethol

Yrgrall Adnoddau y Coleg Cymraeg Cenedlaethol
www.colegcymraeg.ac.uk/cy/

 Coleg
Cymraeg
Cenedlaethol

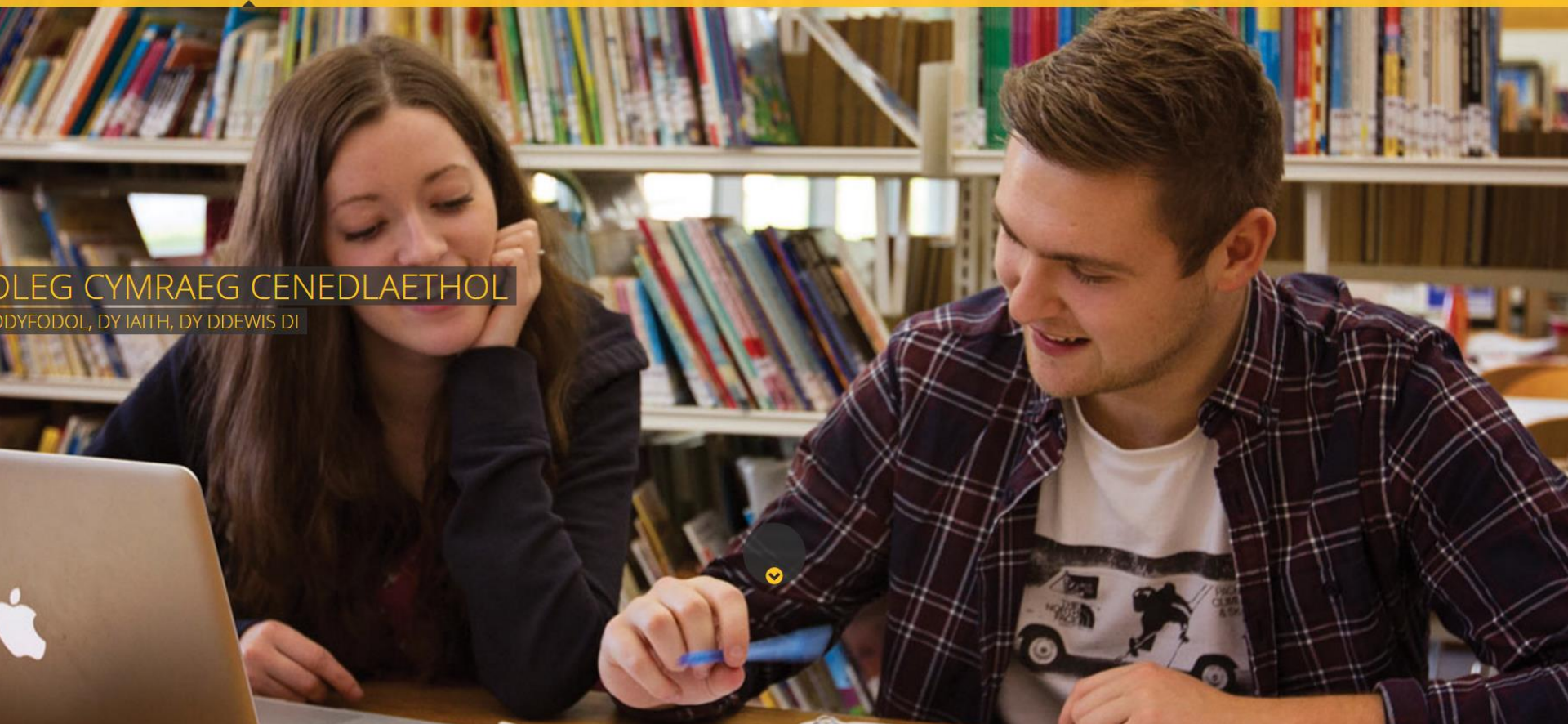
HAFAN ASTUDIO YMAELODI ADNODDAU Y COLEG AMDANOM NI YMCHWIL



ENGLISH

COLEG CYMRAEG CENEDLAETHOL

ADYFODOL, DY IAITH, DY DDEWIS DI



Challenges

- New demand for higher education resources through the medium of Welsh
- Many important scholarly books out of print
- Reprinting on paper is expensive and books will become out of print again
- Welsh Government's Language Strategy lists e-books as an important tool to strengthen infrastructure for Welsh

Solution

E-publishing project: DEChE – Digitization, E-publishing and Electronic Corpus

- Digitizing academic books that are out of print, creating free open access e-books.
- Additional benefits for users. They can:
 - work anywhere
 - search text and create notes
 - listen with Welsh text-to-speech software
- Create Welsh corpus of the e-books' content to support independent research
- Work undertaken by Bangor University's Language Technologies Unit

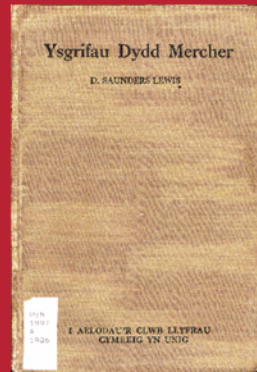
Selecting and Prioritizing Books

- Nomination process arranged by Coleg staff
- Criteria for e-publication:
 - 1) core material for Welsh-medium module
 - 2) on students' reading lists
 - 3) key research material
- 189 titles nominated by 41 different academics spanning over 20 subjects.

Copyright issues

- Permission sought for works still within copyright period
- Small fee offered
- Publishers paid to help trace copyright holders
- Goodwill towards project within publishing houses. All but one of publisher consented to work with project
- The project had no intention to undermine the Welsh publishing industry, if a nominated book is commercially viable

Digitization Process



Original book



EPUB

MOBI

PDF

<https://llyfrgell.porth.ac.uk/>

Scan to TIFF files

OmniPage

clean

auto-proofread

convert to EPUB

Programme to tidy
html markup

MOBI and PDF files
from final EPUB

Calibre

Test on e-readers and
second proofread

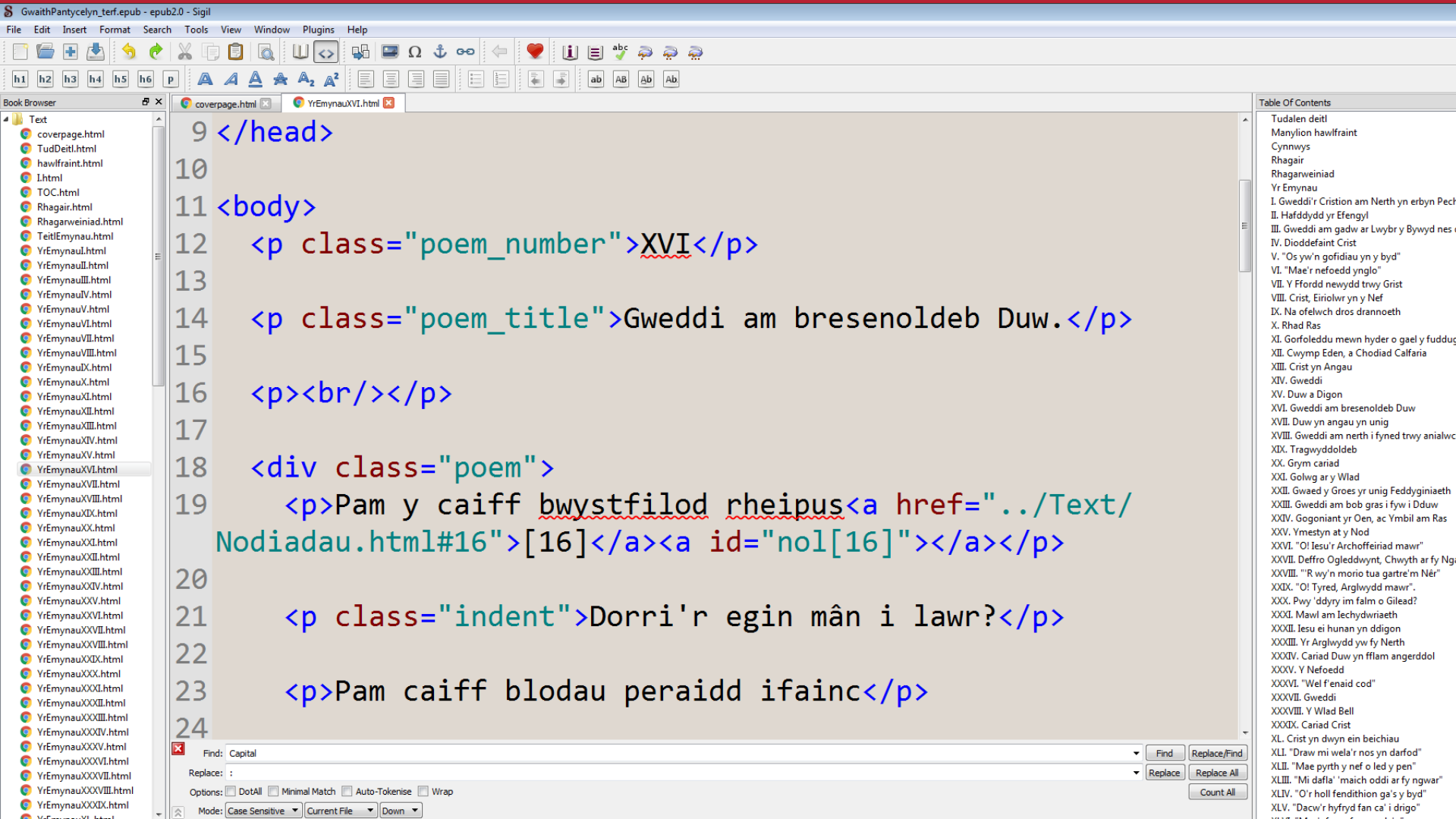
metadata, cover etc
human proofreading

Sigil

Close

Sigil

e-book editing software



Sigil e-book editing software interface showing the HTML code for a book cover page. The code includes a head section, a body section with a poem number 'XVI', a poem title 'Gweddi am bresenoldeb Duw.', a line break, a poem div containing a paragraph about 'Pam y caiff bwystfilod rheipus' and a link to 'Nodiadau.html#16', an indented paragraph 'Dorri'r egin mân i lawr?', and another paragraph 'Pam caiff blodau peraid ifainc'.

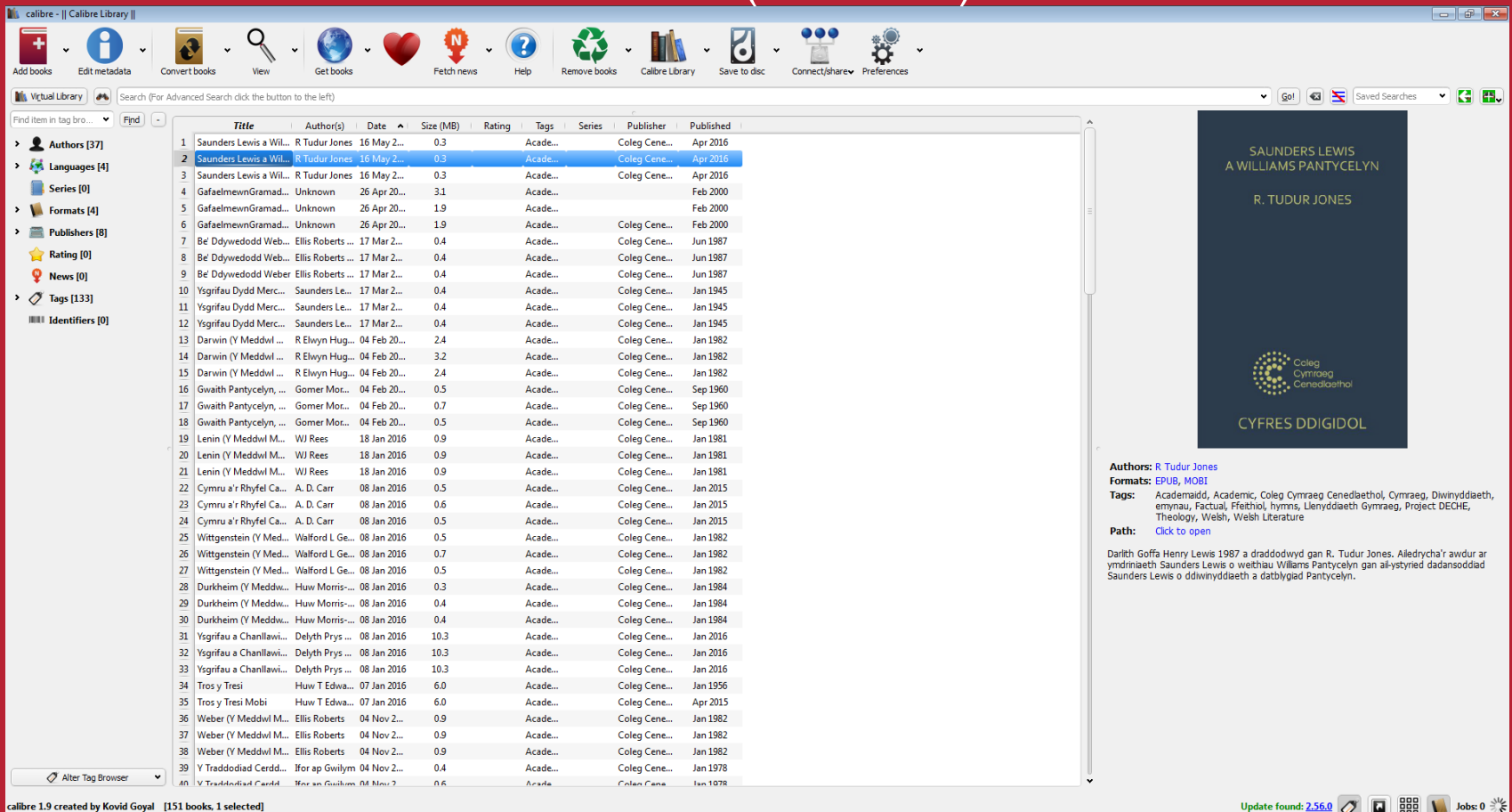
Table Of Contents:

- Tudalen deitl
- Manylion hawffraint
- Cynnwys
- Rhagair
- Rhagarweiniad
- Yr Emynu
- I. Gweddi'r Cristion am Nerth yn erbyn Pech
- II. Hafddydd yr Efengyl
- III. Gweddi am gadw ar Lwybr y Bywyd nes
- IV. Dioddefaint Crist
- V. "Os yw'n gofidiau yn y byd"
- VI. "Mae" nefoedd ynglo"
- VII. Y Ffordd newydd trwy Grist
- VIII. Crist, Eiriolwr yn y Nef
- IX. Na ofelwch dros drannoeth
- X. Rhad Ras
- XI. Gorfoleddu mewn hyder o gael y fuddug
- XII. Cwmp Eden, a Chodiad Calafia
- XIII. Crist yn Angau
- XIV. Gweddi
- XV. Duw a Digion
- XVI. Gweddi am bresenoldeb Duw
- XVII. Duw yn angau yn unig
- XVIII. Gweddi am nerth i fyned trwy analwch
- XIX. Tragwyddoldeb
- XX. Grym cariad
- XXI. Golwg ar y Wlad
- XXII. Gwaed y Groes yr unig Feddyginiaeth
- XXIII. Gweddi am bob gras i fyw i Dduw
- XXIV. Gogoniant yr Oen, ac Ymbil am Ras
- XXV. Ymestyn at y Nod
- XXVI. "O! Iesu'r Archofeiriad mawr"
- XXVII. Deffro Ogledwynt, Chwyth ar fy Ng
- XXVIII. "R wy'n morio tua gartre'm Nêr"
- XXIX. "O! Tyred, Arglwydd mawr".
- XXX. Pwy 'ddryr im falm o Gilead?
- XXXI. Mawl am Iechydwr iaeth
- XXXII. Iesu ei hunan yn ddigon
- XXXIII. Yr Arglwydd yw fy Nerth
- XXXIV. Cariad Duw yn fflam angerddol
- XXXV. Y Nefoedd
- XXXVI. "Wel f'enaid cod"
- XXXVII. Gweddi
- XXXVIII. Y Wlad Bell
- XXXIX. Cariad Crist
- XL. Crist yn dwyn ein beichiau
- XLI. "Draw mi wela'r nos yn darford"
- XLII. "Mae pyth y nef o led y pen"
- XLIII. "Mi dafia" 'maich oddi ar fy ngwar"
- XLIV. "O'r holl fendithion ge's y byd"
- XLV. "Dacw'r hyfryd fan ca' i drigo"
- XLVI. "Mae'r ffrwyth yn ymuno"

Line: 1, Col: 3 210%

Calibre

Open source e-book management program to convert EPUB file to MOBI format (for Kindle) and PDF format



calibre - || Calibre Library ||

Find item in tag browser... Find

	Title	Author(s)	Date	Size (MB)	Rating	Tags	Series	Publisher	Published
1	Saunders Lewis a Williams Pantycelyn	R Tudur Jones	16 May 2016	0.3		Academaidd		Coleg Cymraeg Cenedlaethol	Apr 2016
2	Saunders Lewis a Williams Pantycelyn	R Tudur Jones	16 May 2016	0.3		Academaidd		Coleg Cymraeg Cenedlaethol	Apr 2016
3	Saunders Lewis a Williams Pantycelyn	R Tudur Jones	16 May 2016	0.3		Academaidd		Coleg Cymraeg Cenedlaethol	Apr 2016
4	Gafaelmewn Gramadeg	Unknown	26 Apr 2000	3.1		Academaidd			Feb 2000
5	Gafaelmewn Gramadeg	Unknown	26 Apr 2000	1.9		Academaidd			Feb 2000
6	Gafaelmewn Gramadeg	Unknown	26 Apr 2000	1.9		Academaidd		Coleg Cymraeg Cenedlaethol	Feb 2000
7	Be' Ddyweddod Web...	Ellis Roberts ...	17 Mar 2000	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jun 1987
8	Be' Ddyweddod Web...	Ellis Roberts ...	17 Mar 2000	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jun 1987
9	Be' Ddyweddod Web...	Ellis Roberts ...	17 Mar 2000	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jun 1987
10	Ysgrifau Dydd Merc...	Saunders Le...	17 Mar 2000	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1945
11	Ysgrifau Dydd Merc...	Saunders Le...	17 Mar 2000	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1945
12	Ysgrifau Dydd Merc...	Saunders Le...	17 Mar 2000	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1945
13	Darwin (Y Meddwl M...)	R Elwyn Hug...	04 Feb 2000	2.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
14	Darwin (Y Meddwl M...)	R Elwyn Hug...	04 Feb 2000	3.2		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
15	Darwin (Y Meddwl M...)	R Elwyn Hug...	04 Feb 2000	2.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
16	Gwaith Pantycelyn...	Gomer Mor...	04 Feb 2000	0.5		Academaidd		Coleg Cymraeg Cenedlaethol	Sep 1960
17	Gwaith Pantycelyn...	Gomer Mor...	04 Feb 2000	0.7		Academaidd		Coleg Cymraeg Cenedlaethol	Sep 1960
18	Gwaith Pantycelyn...	Gomer Mor...	04 Feb 2000	0.5		Academaidd		Coleg Cymraeg Cenedlaethol	Sep 1960
19	Lenin (Y Meddwl M...)	WJ Rees	18 Jan 2016	0.9		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1981
20	Lenin (Y Meddwl M...)	WJ Rees	18 Jan 2016	0.9		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1981
21	Lenin (Y Meddwl M...)	WJ Rees	18 Jan 2016	0.9		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1981
22	Cymru a'r Rhyfel Ca...	A. D. Carr	08 Jan 2016	0.5		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 2015
23	Cymru a'r Rhyfel Ca...	A. D. Carr	08 Jan 2016	0.6		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 2015
24	Cymru a'r Rhyfel Ca...	A. D. Carr	08 Jan 2016	0.5		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 2015
25	Wittgenstein (Y Meddwl M...)	Walford L. Ge...	08 Jan 2016	0.5		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
26	Wittgenstein (Y Meddwl M...)	Walford L. Ge...	08 Jan 2016	0.7		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
27	Wittgenstein (Y Meddwl M...)	Walford L. Ge...	08 Jan 2016	0.5		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
28	Durkheim (Y Meddwl M...)	Huw Morris...	08 Jan 2016	0.3		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1984
29	Durkheim (Y Meddwl M...)	Huw Morris...	08 Jan 2016	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1984
30	Durkheim (Y Meddwl M...)	Huw Morris...	08 Jan 2016	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1984
31	Ysgrifau a Chanllaw...	Delyth Prys ...	08 Jan 2016	10.3		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 2016
32	Ysgrifau a Chanllaw...	Delyth Prys ...	08 Jan 2016	10.3		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 2016
33	Ysgrifau a Chanllaw...	Delyth Prys ...	08 Jan 2016	10.3		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 2016
34	Tros y Tres	Huw T. Edwa...	07 Jan 2016	6.0		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1956
35	Tros y Tres	Huw T. Edwa...	07 Jan 2016	6.0		Academaidd		Coleg Cymraeg Cenedlaethol	Apr 2015
36	Weber (Y Meddwl M...)	Ellis Roberts	04 Nov 2000	0.9		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
37	Weber (Y Meddwl M...)	Ellis Roberts	04 Nov 2000	0.9		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
38	Weber (Y Meddwl M...)	Ellis Roberts	04 Nov 2000	0.9		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1982
39	Y Traddodiad Cerdd...	Ifor ap Gwilym	04 Nov 2000	0.4		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1978
40	Y Traddodiad Cerdd...	Ifor ap Gwilym	04 Nov 2000	0.6		Academaidd		Coleg Cymraeg Cenedlaethol	Jan 1978

calibre 1.9 created by Kovid Goyal [151 books, 1 selected]

Update found: 2.56.0

Jobs: 0

Llyfrgell Adnoddau (Resource Library)

<https://llyfrgell.porth.ac.uk>

Llyfrgell Adnoddau y Coleg Cymraeg Cenedlaethol

Hafan Llyfrgell

Mewngofnodi

Geiriadur Prifysgol Cymru
1 mis yn ôl
Mae ap newydd (Apple neu Android) ar gael yn rhad ac am ddim sy'n rhoi mynediad llawn i Geiriadur Prifysgol Cymru (GPC), y geiriadur hanesyddol Cymraeg sy'n cyfateb i Oxford English Dictionary. Mae hwn yn waith enfawr o dros 4,000 o dudalennau a thros wyth miliwn o eiriau i gyd. Mae'n anhepgorol i bawb sy'n astudio! Gymraeg a'i llenyddiaeth gyfoethog, ond mae hefyd yn ddefnyddiol iawn mewn llawer o feysydd eraill. Ceir diffiniadau yn Gymraeg yn ogystal â chyfystyron Saesneg, a dyfyniadau yn dangos defnydd y gair dros y canrifoedd. Rhestrir cyfyniadau cyffredin sy'n cynnwys y gair, gan gynnwys nifer o dermau technegol. Mantais fawr y fersiwn electronig yw bod modd chwilio am air Saesneg yn y Geiriadur, peth sydd heb gael ei wneud yn y Geiriadur, peth sydd heb gael ei wneud yn y Geiriadur, peth sydd heb gael ei wneud yn y Geiriadur.

ceffyl, cel
1/73
[H. Wydd. capai' ceffyl gwaith', celi. gathre. ll. capliu, Gal. a Llad. caballus].
eg. ll. ceffylau.
Anafal pedwarcanol cryf a chylfym a chanddo feng a cyrffon, a ddofwyd gan ddyn i'w farchogaeth a hefyd i gludo beichiau ac i dynnu cerbydau, &c., march a ddisbuddwyd yn eiddo.
horse, nag. hobby.
14g. MS. 119.15. y ceffylau a gwreiddio gronud.
14g. MS. 56. Fal ceffyl trymugl tremgrog.
1567. TN. 3930. a' chysgyle (-) meirch, a' sauridry.
1588. 1. Br. xviii. 5. y ceffylau a'i mudo.
1732-3. J. Owen. GB. 84. a' Car o flaen y Cel.
Ardur. J. Owen. GB.
chp. J. Owen. Owen (Nerfen Angered). Goleg ar y ym (Deiau. 1732-3).

Tystysgrif Sgiliau Iaith Coleg Cymraeg Cenedlaethol
7 mis yn ôl
Myfyrwyr yn siarad am fanteision y Dystysgrif Sgiliau Iaith

Ap Chwilydd Cyrsiau'r Coleg Cymraeg
1 blynedd yn ôl
Chwilio am gars cyflwyn Cymraeg mewn prifysgol yng Nghymru? Lawrwytho Ap Chwilydd Cyrsiau'r Coleg Cymraeg! Mae'r Ap ar gael ar declynnau iPad, iPhone ac Android.

E-lyfrau

Saunders Lewis a Williams Pantycelyn – R. Tudur Jones
2 wythnos yn ôl
Darlith Goffa Henry Lewis 1987 a draddodwyd gan R. Tudur Jones. Edrychfa'r awdur ar ymdriniadau Saunders Lewis a Williams Pantycelyn.

Gafael Mewn Gramadeg – David Thorne
2 wythnos yn ôl
Astudiaeth gyfoes o ramadeg y Gymraeg sy'n amcanu i esbonio ei hanfodion. Y mae'r gyfrol yn cynnwys trafodaeth ar gywair ac amrywiaeth eang o arddulliau llaf...

Ysgrifiad Dydd Mercher – Saunders Lewis
1 mis yn ôl
Casgliad o adolygiadau ac ysgrifiad cofannol gan Saunders Lewis a gyhoeddwyd yn wreiddiol yn Y Faner rhwng 1939 ac 1945. Noda Saunders Lewis yn y...

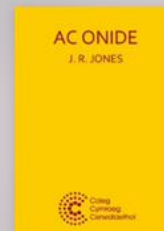
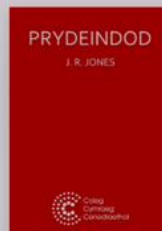
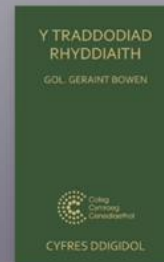
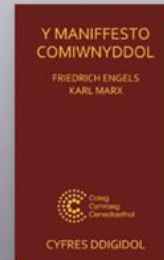
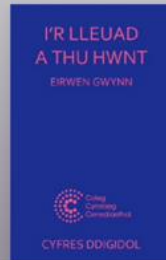
Gwaith Pantycelyn: Detholiad – Gomer M. Roberts (gol.)
3 mis yn ôl
Detholiad o gant o emynau William Williams Pantycelyn ynghyd â dwy farwnad, dyfyniadau o gerddi hirach a...

15:53
16/05/2016



All Books ▾

Select



PRIFYSGOL
BANGOR
UNIVERSITY

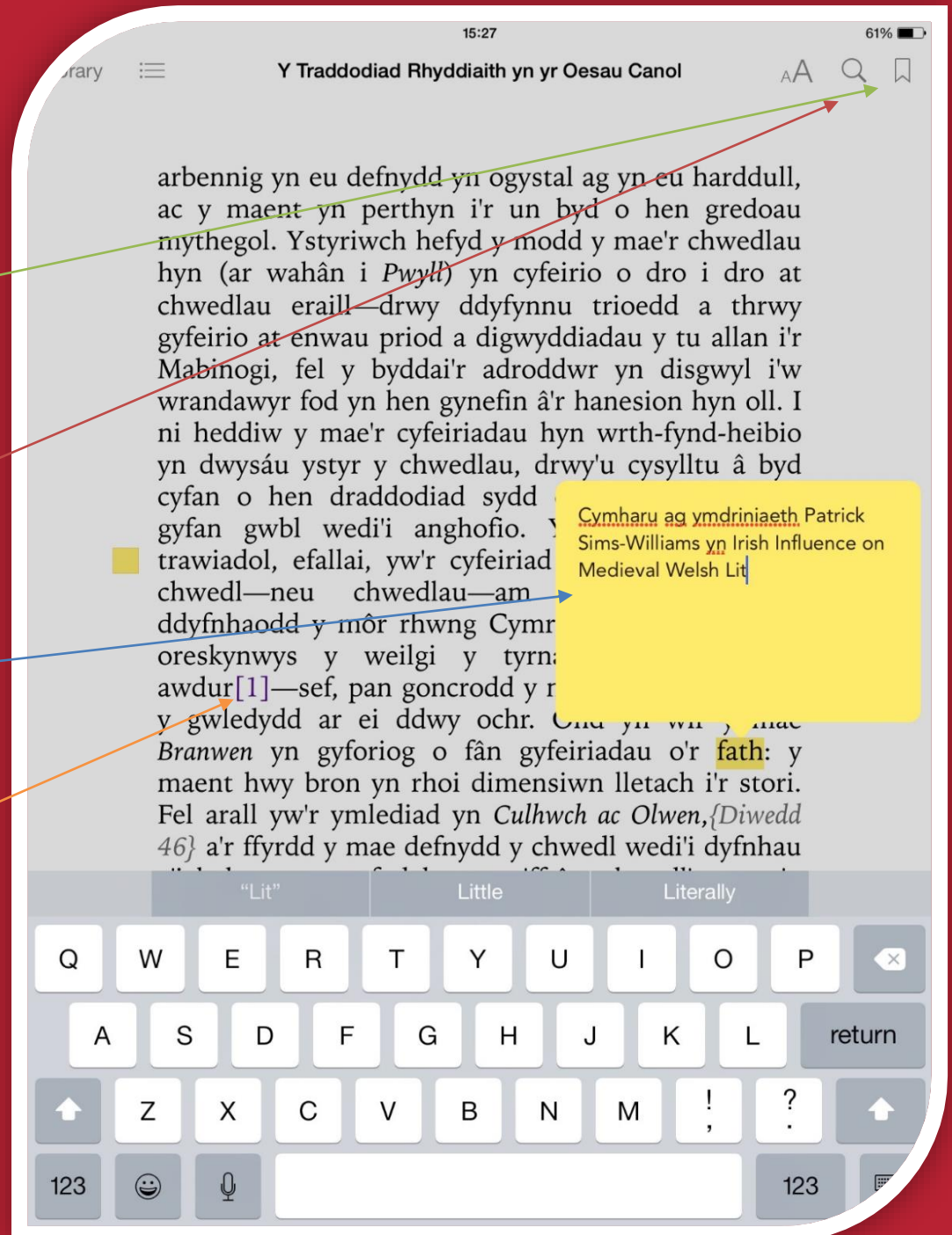
Within EPUB and MOBI:

Bookmark

Search text

Write notes

Toggle between
endnotes and
main text





Mynhodd ef enghyd y mair amiaid o daigon ar gywydd ac enghyd ac ar benillion rhydd yn Cymraeg a Saesneg.

Mae'r cyffro hwn i'w weld ar ei fwyaf cymhleth yn y disgrifiad a **rydd** yr awdur i ni o Genfigen.

M wedd

30

610 results in 44 documents

A

[8

lla

W

gy

fo

dr

A

hwn, "Dychymmygu llun neu ddelw o Dduw yn ein meddyliau; canys delwaddoliad ysbrydol yw hyny".

Nid yw'r darluniau a'r delweddau Beiblaidd chwaith yn **rhydd** o berygl oni ofelir eu cadw yn eu priod le a rhoi iddynt eu priod swydd trwy eu trin yng ngolau'r athrawiaeth.

Mae Charles am weld llenorion yn ysgrifennu: {Diwedd 347} mewn **rhydd-iaith** bur, syml, a chref ar wahanol destynau sylweddol, pwysig, ac adeiladol, a barant eangiad meddwl, helaethiad gwybodaeth, gynnyrchiaid myfyrdod, a hyfforddiant gwirioneddol i'r genedl.

Diau petai'r ysgrifenywyr wedi glynu wrth y "**rhydd-iaith** bur, syml a chref" y soniodd David Charles amdani, y byddai gwell

n?

u

ylw

na

caol

IE

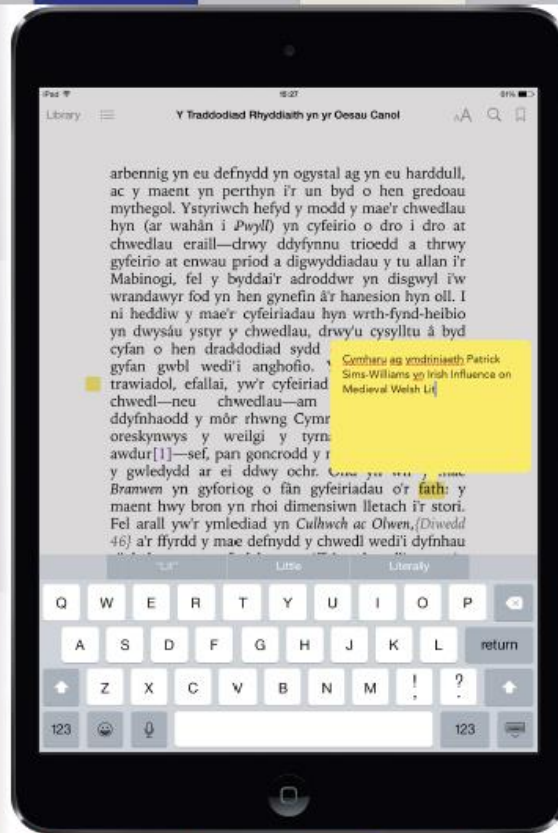
irwydd

CYFRES DDIGIDOL

llyfrgell.porth.ac.uk



Tweets



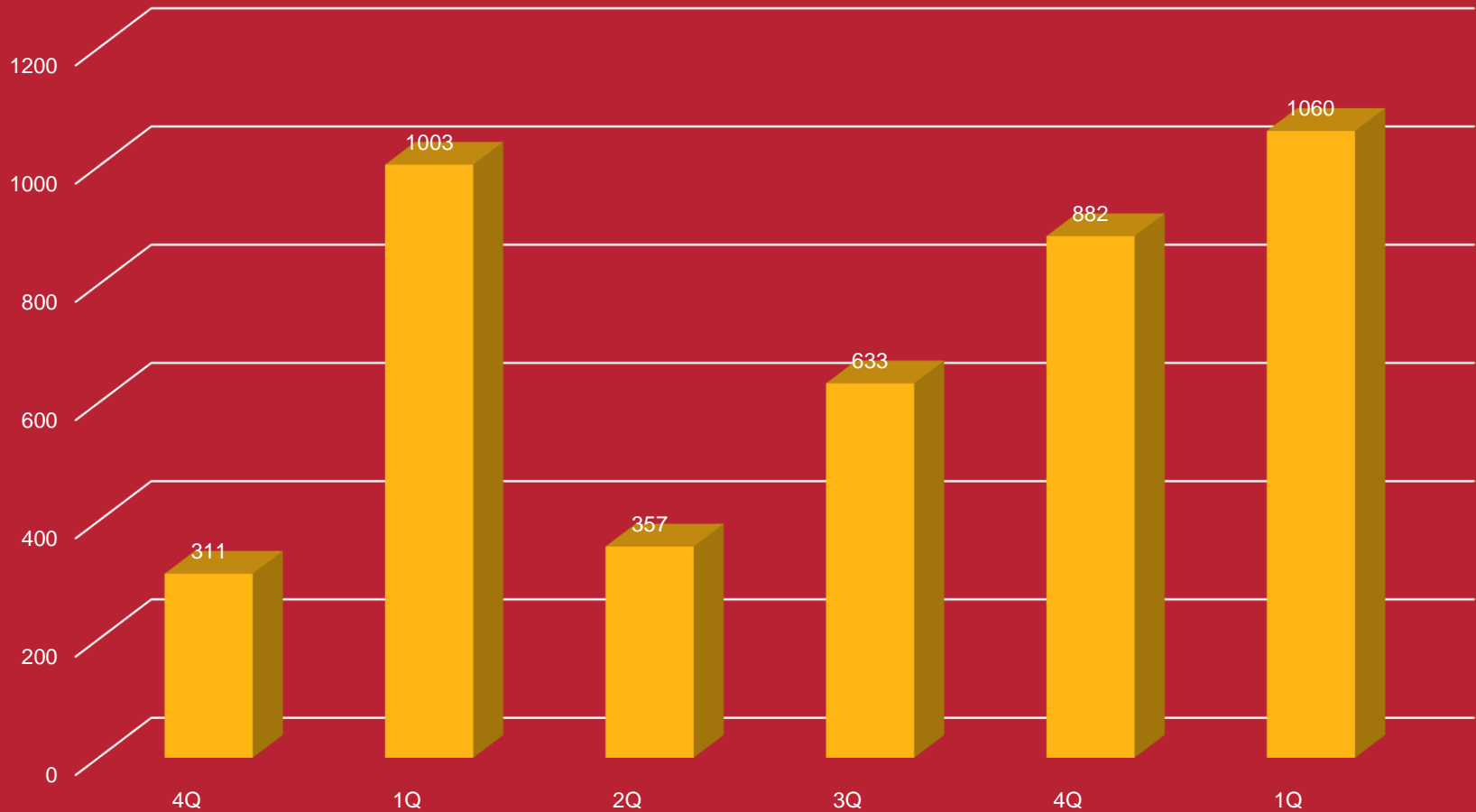
- E-lyfrau
- EPUB
- Mobi (Kindle)
- PDF

AM DDIM!

CLASURON YSGOLHEIGAIDD
AR EICH FFÔN, PC NEU DABLED

Total Unique Hits

(per quarter, Oct 2014 to March 2016)



Conclusions

- Successful method of republishing academic books
- Weakness in uptake and visibility being addressed
- Important works sustainable in long term
- Stimulate original scholarly discourse
- Status in digital media for minority language
- Marketing tool for Coleg to wider audience
- Language corpus as secondary benefit

Diolch yn fawr!



Y Coleg ar y Bryn