Corpws Cenedlaethol Cymraeg Cyfoes

National Corpus of Contemporary Welsh

Cynnwys

T2: Newyddion

P4: Diweddariadau ar Becynnau Gwaith (WP)

P6: Cwrdd â'r tîm

P7: Cysylltu â ni

Cyfarchion gan y Prif Ymchwilydd

Croeso i rifyn mis Medi cylchlythyr CorCenCC! Mae'r haf (yn anffodus) drosodd ac rydym ar drothwy'r flwyddyn academaidd newydd, a 24 mis olaf prosiect CorCenCC. Mae misoedd yr haf wedi cynnwys adegau dwys o gasglu data gan y tîm, yn ogystal â rhywfaint o ledaenu gwaith y prosiect mewn ystod o gynadleddau academaidd. Yn wythnos gyntaf mis Medi, daeth holl aelodau'r tîm CPT (y Prif Ymchwilydd, y Cyd-Ymchwilwyr a'r Cynorthwywyr Ymchwil) ynghyd, gan deithio o Fangor, Caerhirfryn ac Abertawe i Gaerdydd ar gyfer ein cyfarfod chwemisol rheolaidd. Cyfarfod amlswyddogaethol oedd hwn a roddodd amser i'r tîm craidd gwrdd wyneb yn wyneb, pwyso a mesur y cynnydd a wnaed hyd yn hyn a myfyrio'n gadarnhaol arno, a blaenoriaethu a chynllunio misoedd nesaf y prosiect. Roedd yn gyfarfod cynhyrchiol iawn ac, fel ag erioed, roedd yn wych dal i fyny gyda phawb eto! Yn rhifynnau nesaf y cylchlythyr, byddwn yn darparu diweddariadau penodol ar y Pecynnau Gwaith i chi (gan ddechrau gyda Phecynnau Gwaith 1, 3 a 4). Byddwn hefyd yn darparu manylion am ein gweledigaeth a'n cynlluniau ar gyfer 24 mis olaf y prosiect ac, fel arfer, rhagor o fanylion am sut y gallwch chi gymryd mwy o ran!

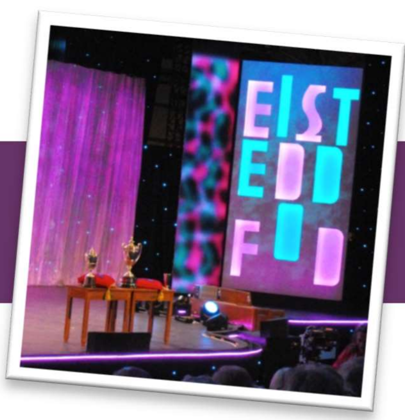
Mwynhewch y cylchlythyr! Dr Dawn Knight

Siarad Cymraeg?

Bydd croeso i unrhyw un sy'n siarad Cymraeg alw yn Nhŷ'r Gwrhyd, Pontardawe ar 17 Tachwedd rhwng 13:00 a 15:00 a rhoi cynnig ar fod yn feirniad ffilm am y prynhawn! Byddwn yn dangos clipiau byr o wahanol fideos ac yn rhoi cyfle i chi a'ch ffrindiau ddweud wrthon ni beth ŷch chi'n meddwl amdanyn nhw. Bydd cyfle gan y rhai sy'n mynd i'r digwyddiad (sy'n rhan o raglen Abertawe ar gyfer gŵyl flynyddol '*Being Human*') i glywed mwy am brosiect CorCenCC... yn ogystal â dysgu am sut i lawrlwytho a defnyddio'r ap. Dewch draw, rhowch eich Cymraeg i ni a mynnwch gwrdd â Steve Morris a Tess Fitzpatrick o'r tîm i wybod mwy am y prosiect.

+ Newyddion a digwyddiadau

Ym mis Awst, teithiodd tri o gynorthwywyr ymchwil CorCenCC i Ynys Môn i ymweld â'r Eisteddfod Genedlaethol. Fel un o wyliau pwysicaf Cymru, lle mae pob perfformiad, cystadleuaeth a seremoni'n digwydd trwy gyfrwng y Gymraeg, roedd hwn yn gyfle ardderchog i ni gwrdd â nifer fawr o siaradwyr Cymraeg a'u hannog nhw i gyfrannu at ein prosiect. Bwriad prosiect CorCenCC yw adlewyrchu'r ystod eang o wahanol ffyrdd y mae'r Gymraeg yn cael ei defnyddio heddiw. Felly mae'n bwysig iawn cynnwys elfennau sy'n unigryw i ddiwylliant Cymru, megis pobl yn defnyddio'r Gymraeg yn yr Eisteddfod.
Casglon ni enghreifftiau o bobl yn beirniadu cystadlaethau, yn rhoi darlithoedd neu gyflwyniadau, ac – wrth gwrs – yn mwynhau dros beint, paned, neu hufen iâ! Wedi dweud hynny, doedd hi ddim yn dywydd hufen iâ bob dydd, yn anffodus. Yn sgil y glaw trwm a'r gwyntoedd cryf, roedd hi'n heriol dod o hyd i lefydd addas i recordio … ac weithiau unrhyw bobl i'w recordio! Ond gwenodd yr haul arnon ni hefyd, ac roedd hi'n bleser cwrdd â siaradwyr Cymraeg a oedd wedi dod o bob cwr o Gymru i ddathlu eu hiaith a'u diwylliant. Diolch yn fawr iawn i bawb a gytunodd i gael ei recordio – mae'ch cyfraniadau wedi'n helpu ni i sicrhau bod y corpws wir yn cynnwys tipyn o bopeth!



Ond nid gŵyl ar gyfer siaradwyr rhugl yn unig yw'r Eisteddfod. Un o'r pethau gwych amdani yw'r ffaith ei bod yn cynnig cyfle rhagorol i bobl sy'n dysgu'r Gymraeg. Gyda chymaint o Gymraeg o'u cwmpas, mae'r Eisteddfod yn lle delfrydol i ddysgwyr glywed ac ymarfer Cymraeg, ac mae cyfres o gystadlaethau arbennig ar gael ar gyfer dysgwyr o bob lefel.

Mae'r Eisteddfod yn ŵyl deithiol, ac yn 2018 bydd hi'n dod i 'gartref' CorCenCC, sef Caerdydd, felly gallwch chi fod yn sicr y bydd criw ohonon ni yno yn cynrychioli'r prosiect! Erbyn hynny, bydd rhan o'r corpws wedi'i hadeiladu a bydd modd i chi chwarae â rhai o'r teclynnau y byddwn ni wedi'u datblygu – felly dewch yn llu i gwrdd â ni! https://eisteddfod.cymru/

BAAL

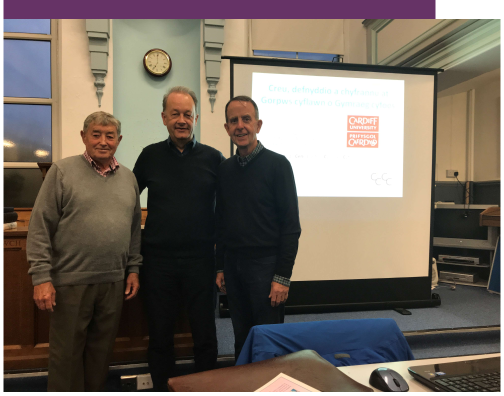
Cynhadledd BAAL 2017, Prifysgol Leeds 31/8 - 2/9

Cymdeithas broffesiynol gydag ychydig o dan fil o aelodau rhyngwladol yw BAAL. Mae'n darparu fforwm ar gyfer pobl sydd â diddordeb mewn iaith ac ieithyddiaeth gymhwysol. Am ragor o wybodaeth ewch i: www.baal.org.uk

Roedd yn dda gan Dîm Rheoli CorCenCC (Dawn Knight, Steve Morris a Tess Fitzpatrick) gael y cyfle i gyflwyno poster a sesiwn yng nghynhadledd BAAL 2017 yn Leeds ddechrau mis Medi. Amlinellodd ein poster y gwaith yr oedd Steve a Tess wedi'i wneud o'r blaen ar greu rhestri geirfa pedagogaidd yn Gymraeg gan ddefnyddio technegau a ysbrydolwyd gan Français Fondamental (Gougenheim et al, 1964; Morris 2011) a chan ymchwil cysylltu geiriau. Mae CorCenCC yn cynnig dull o gasglu rhestri geiriau ar sail amlder geiriau, ac mi fydd cymharu hyn gyda'r rhestri cynharach yn ysgogi cwestiynau ymchwil newydd! Cyflwynodd Dawn a Steve ein papur am nodau'r prosiect, a'r cynnydd hyd yn hyn. Roedd y gynulleidfa yn cynnwys ieithyddion cymhwysol sy'n gweithio ar ieithoedd prin eu hadnoddau yn ogystal â'r rhai sy'n ymddiddori mewn ieithyddiaeth corpws, a chawsom ein calonogi gan y nifer o sylwadau oedd yn ymwneud â chapásiti CorCenCC i weithredu fel glasbrint ar gyfer gwaith corpws mewn ieithoedd eraill.

CorCenCC ar daith

27/9



Un elfen bwysig o brosiect CorCenCC yw 'ymgysylltu' ac ar noson wlyb iawn ym mis Medi, aeth Steve Morris ar wahoddiad i siarad â Chymdeithas y Llan a'r Bryn yn Llangennech, Sir Gâr. Roedd diddordeb mawr gan y gynulleidfa yn y prosiect a thipyn o drafodaeth am gynnwys 'iaith sathredig' yn y corpws a beth yn union yw 'Cymraeg cywir'. Cytunodd nifer o'r bobl a aeth i'r sesiwn gyfrannu eu Cymraeg i'r prosiect (dyma ardal y "*Shwt ife heddi?*" enwog) a byddwn yn recordio sesiwn gyfan o raglen 2017-18 y Gymdeithas yn nes ymlaen yn y flwyddyn. Wrth i'r prosiect dyfu a datblygu, ein bwriad yw mynd â CorCenCC '*ar daith*' yn fwy aml er mwyn ymygysylltu â'r gymuned Gymraeg ehangach. Os oes unrhyw awgrymiadau gyda chi, cysylltwch ar bob cyfrif!

Llun: Glendon Davies (Trysorydd), Steve Morris, Colin Lee (Trefnydd Rhaglen)

Diweddariadau ar Becynnau Gwaith (WP1/3/4)

Fel y byddwch chi wedi gweld yn ein cylchlythyrau blaenorol, byddwn ni'n casglu enghreifftiau o Gymraeg llafar, ysgrifenedig ac electronig i'w cynnwys yn y corpws. Ein prif flaenoriaeth ar hyn o bryd yw casglu'r enghreifftiau llafar, gan fod y rhain yn gofyn am lefel ychwanegol o brosesu – sef, trawsgrifio. Mae ein tîm o drawsgrifwyr yn gweithio'n galed iawn, ac rydyn ni wir yn gwerthfawrogi eu cymorth o ran paratoi cynnwys y corpws. Diolch o galon!

Os ydych chi am ymuno â'n tîm o drawsgrifwyr, anfonwch e-bost at trawsgrifio@corcencc.org i gael manylion pellach.

Wrth i'r trawsgrifwyr weithio ar greu cofnod ysgrifenedig (teipiedig) o gynnwys y recordiadau, mae'r cynorthwywyr ymchwil wedi bod yn brysur yn casglu rhagor o recordiadau! Ers ein diweddariad diwethaf, rydyn ni wedi recordio siaradwyr yng Ngwynedd, Ceredigion a De Cymru, ac yn Eisteddfod yr Urdd a'r Eisteddfod Genedlaethol. Mae'n hyfryd cael dod i adnabod gwahanol ardaloedd, ac fe fyddwn ni'n ymweld â phob sir yng Nghymru erbyn diwedd y cyfnod casglu data. Anfonwch e-bost at corcencc@caerdydd.ac.uk os ydych chi am wybod pryd byddwn ni yn eich ardal chi. Hoffen ni gwrdd â chymaint ohonoch chi â phosib, er mwyn wir adlewyrchu sut mae defnydd yr iaith Gymraeg yn amrywio o ardal i ardal, o gyd-destun i gyd-destun, ac o berson i berson.

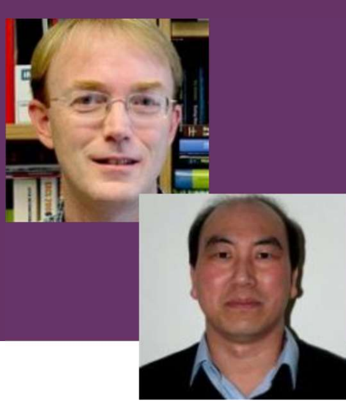
WP1:

Casglu, trawsgrifio ac anonymeiddio'r data (arweinydd: Steve Morris)

Rydyn ni hefyd wedi dechrau casglu enghreifftiau o iaith y cyfryngau, drwy recordio darnau o raglenni teledu S4C. Cyn bo hir, rydyn ni'n gobeithio dechrau recordio darnau o raglenni BBC Radio Cymru hefyd. Mae'r Gymraeg yn ffodus o gael sianeli radio a theledu arbennig, ac mae CorCenCC yn ffodus iawn bod S4C a'r BBC wedi cytuno i gydweithio â ni, i'n helpu ni i greu 'ciplun' o Gymraeg cyfoes yng Nghymru.

WP3:

Datblygu tagiwr semantig i'r Gymraeg a thagio ystyr yr holl ddata (arweinydd: Paul Rayson)



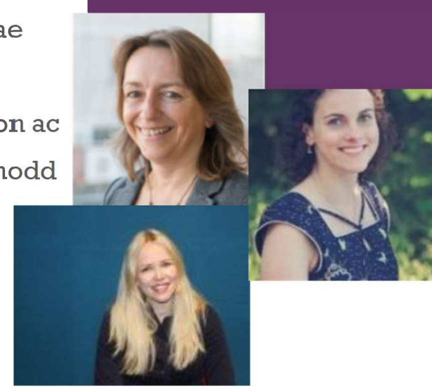
Ym Mhecyn Gwaith 3, rydym wedi bod yn parhau â'r ymchwil ar ddatblygu tagiwr semantig effeithlon i'r Gymraeg, gan gynnwys adeiladu lecsicon semantig mwy er mwyn cynyddu'r nifer o eiriau y gall y tagiwr eu cwmpasu, ac er mwyn cyflwyno dulliau mwy effeithiol ar gyfer nodi ystyron cywir geiriau. Rydym hefyd wedi bod wrthi'n ymdrechu i sicrhau bod y tagiwr semantig yn hygyrch a bod defnyddwyr yn gallu cael ati yn hwylus. Yn ysgol haf UCREL yng Nghaerhirfryn ym mis Mehefin 2017, cafodd y fersiwn gyfredol o'r tagiwr semantig ei chyflwyno a'i hyrwyddo i ieithyddion corpws ac ymchwilwyr prosesu iaith naturiol, fel rhan o system tagio semantig amlieithog UCREL. Rydym yn darparu mynediad i'r system drwy wasanaeth API ar y we a thrwy rhaglen bwrdd gwaith (mae Rhyngwyneb Defnyddiwr Graffigol ar gyfer gwasanaeth gwe y tagiwr Cymraeg ar gael i'w lawrlwytho o <http://ucrel.lancs.ac.uk/usas/gui/>). Er mai prototeip ydyw ar hyn o bryd,

mae'n offeryn defnyddiol ar gyfer echdynnu gwybodaeth semantig o ddata Cymraeg a'i dadansoddi.

Mae'r camau mawr nesaf yn cynnwys 1) integreiddio'r tagiwr rhan ymadrodd newydd i'r Gymraeg (CyTag) sydd dan ddatblygiad gan Brosiect CorCenCC, ac sy'n defnyddio rhestr fanwl o dagiau rhan ymadrodd sydd wedi'i dylunio'n ddiweddar; 2) yn seiliedig ar y corpws safon aur sy'n cael ei anodi â llaw ar hyn o bryd, byddwn yn gwerthuso, yn cloriannu ac yn integreiddio mwy o ddulliau o wahaniaethu rhwng ystyron geiriau ar sail cyd-destun; 3) bydd mwy o eiriau ac ymadroddion aml-eiriau Cymraeg yn cael eu dosbarthu o ran ystyr semantig ac yn cael eu hymgorffori yn y lecsicon semantig trwy ganfod geiriau ac ymadroddion cyfystyr. Bydd y tagiwr semantig, sydd yn cael ei wella'n ddi-baid, yn cael ei arddangos mewn amryw ddigwyddiadau ieithyddiaeth corpws a digwyddiadau prosesu iaith naturiol yn y flwyddyn i ddod, a byddwn yn ymgymryd â'r gwaith o ddefnyddio'r tagiwr gyda data corpws CorCenCC sydd ar y gweill.

WP4:

Cwmpasu/llunio'r pecyn cymorth pedagogaidd ar-lein (arweinydd: Enlli Thomas)



Ers ein diweddariad diwethaf (mis Mai, rhifyn 11), rydyn ni wedi bod yn cydweithio ag aelodau tîm Pecyn Gwaith 5 (WP5) i ddylunio'r pecyn cymorth pedagogaidd. Tîm WP5 fydd yn rhoi syniadau tîm WP4 ar waith yn dechnolegol i greu'r adnodd ar-lein, felly mae cydweithio agos yn bwysig iawn i sicrhau bod ein cynlluniau'n ymarferol.

Rydyn ni hefyd yn parhau i ymgynghori ag athrawon a dysgwyr Cymraeg, er mwyn seilio dyluniad y pecyn cymorth pedagogaidd ar eu hanghenion. Fodd bynnag, mae llawer o newidiadau ar y gorwel o ran anghenion a disgwyliadau ym myd dysgu Cymraeg, gyda chwricwla newydd yn cael eu datblygu ar gyfer dysgwyr o oedolion ac ysgolion fel ei gilydd. Ym mis Medi, cafodd Enlli Thomas, arweinydd WP4, ei gwahodd i gyflwyno'r cydsyniad sydd y tu ôl i'r pecyn cymorth pedagogaidd i'r grŵp sy'n gyfrifol am ddatblygu maes Ieithoedd, Llythrennedd a Chyfathrebu y cwricwlwm newydd sydd wrthi'n cael ei ddatblygu ar gyfer ysgolion Cymru. Ymhlith yr oddeutu 40 unigolyn o ledled Cymru oedd croestoriad o athrawon o Ysgolion Arloesi a oedd yn cynrychioli sawl Cyfnod dysgu Allweddol, addysg ddwyieithog, cyfrwng Cymraeg a chyfrwng Saesneg, a darpariaeth brif ffrwd ac AAA, ynghyd ag arolygwyr Estyn a swyddogion Is-adran Cwricwlwm ac Asesu Llywodraeth Cymru.

Roedd y drafodaeth yn hynod ddefnyddiol, ac roedd cyfnewid syniadau drwy gydol y bore wedi cyfoethogi ein gweledigaeth o ran sut i lywio'r gwaith wrth symud yn ein blaenau. Mae'r grŵp wedi cytuno i'n gwahodd ni yn ôl maes o law i roi diweddariad iddyn nhw ynghylch ein cynnydd, ac maen nhw'n hapus i roi adborth ar yr hyn yr ydyn ni'n ei ddatblygu. Diolch yn fawr iawn!

A wyddech chi?

Bydd y cyfleusterau pedagogaidd yn cynnwys offer 'codi ymwybyddiaeth ieithyddol' fydd yn gweithredu er mwyn galluogi ac annog dysgwyr (iaith gyntaf ac ail iaith) i ymwneud â phatrymau ieithyddol fydd yn hwyluso'r dysgu. Bydd y rhain yn gymorth o ran creu cwisiau a thasgau i'r dysgwyr sydd wedi'u seilio ar dystiolaeth corpws.

Cwrdd â'r tîm:

Yr Athro Colin Williams, Aelod o Grŵp Ymgynghorol y Prosiect (PAG)

Gwyddonydd Cymdeithasol ydw i, gyda hyfforddiant mewn Daearyddiaeth a Gwleidyddiaeth. Rwyf wedi bod yn Athro Daearyddiaeth ym Mhrifysgol Swydd Stafford, Athro Daearyddiaeth "Fulbright" ym Mhrifysgol Gwladwriaeth Pennsylvania, ac rwy'n parhau i fod yn Athro Dirprwyol mewn Daearyddiaeth ym Mhrifysgol Gorllewin Ontario.

Ar hyn o bryd, rwy'n gweithio yng Ngholeg St Edmund, Caergrawnt lle rwy'n gweithio ar agweddau ar Ailadeiladu Ôl-wrthdaro ac ar Bolisi Amlieithog. Cyn symud i Gaergrawnt yn 2015, roeddwn yn Athro Ymchwil yn Ysgol y Gymraeg, Prifysgol Caerdydd am 21 o flynyddoedd, lle rwy'n parhau i fod yn Athro Anrhydeddus. Yng Nghaerdydd, fy maes diddordeb yw polisi a chynllunio iaith. Ers dros ddegawd, rwyf wedi bod yn cyfuno fy mhrif ddiddordebau ymwchil a'r dysgu yn y brifysgol gyda bod yn aelod o Fwrdd yr Iaith Gymraeg, lle roeddwn yn arbenigo mewn strategaeth, arloesedd, ymyrraeth gan y llywodraeth a pholisi cymharol Ewropeaidd a Chanadaidd. Rwy'n parhau i ddarparu cyngor i Lywodraeth Cymru ar faterion datblygu polisi.

Rwy'n falch o fod yn aelod o brosiect CorCenCC ac rwy'n gwerthfawrogi'r amrywiaeth o ddisgyblaethau a sgiliau arbenigol sydd gan y tîm. Pan gaiff ei harneisio'n llawn, dylai allbynnau cyfunol y tîm ddarparu adnodd ardderchog yn y tymor canolig i gynorthwyo gydag ymdrechion academyddion, gweision cyhoeddus, ymarferwyr a dinasyddion i wella'r anghenion dwyieithog sydd gan Gymru amlieithog.

Yn fy marn i, y rhan o'r gwaith sydd o'r brys mwyaf yw argyhoeddi'r gymdeithas sifil i wneud y gorau o ganlyniadau'r corpws, ond y rhan fwyaf cyffrous yw sylweddoli y bydd CorCenCC yn cyfrannu yn y tymor hir mewn ffyrdd nad ydynt yn bodoli ar hyd o bryd. Yn yr un modd â chronfeydd data pwysig eraill yn y gorffennol, fel Geiriadur Prifysgol Cymru, bydd gweledigaeth sylfaenwyr CorCenCC yn darparu adnodd cyfres-amser rwy'n gobeithio ei weld yn cael ei ddefnyddio mewn o leiaf tair ffordd, sef 1) fel ffynhonnell gwybodaeth a data sy'n adlewyrchu bywiogrwydd ac amrywiaeth y modd y defnyddir y Gymraeg ar hyn o bryd; 2) fel offeryn ar gyfer mireinio polisïau ymyrraeth penodol a rhaglenni mewn arferion addysgol, datblygiad cymunedol, y cyfryngau a datblygiadau technoleg gwybodaeth; a 3) fel llwyfan sylfaenol ar gyfer hyrwyddo'r Gymraeg fel iaith cyflogaeth, masnach a bywyd proffesiynol, er mwyn gwireddu'r potensial llawn o weithio mewn cymdeithas ddwyieithog o fewn cymdeithas amlieithog.

> 'Pan gaiff ei harneisio'n llawn, dylai allbynnau cyfunol y tîm ddarparu adnodd ardderchog i gynorthwyo gydag ymdrechion academyddion, gweision cyhoeddus, ymarferwyr a dinasyddion i wella'r anghenion dwyieithog sydd gan Gymru.'
>
> \- Yr Athro Colin Williams.

Fy rôl i yn CorCenCC yw cynghori ar weithredu strategol ac ar gysylltiad y prosiect ag asiantaethau llywodraethol yng Nghymru a'r tu hwnt.

+ Ac yn olaf…

Fel nodwyd yn ein cylchlythyr diwethaf, mae Gareth Watkins yn ymadael â ni ddiwedd mis Medi. Hoffen ni i gyd ddymuno'n dda i ti, Gareth, wrth i ti ddilyn llwybr newydd yn Yr Wyddgrug a diolch i ti eto am dy waith a'th gyfeillgarwch dros y 19 mis diwethaf ers i'r prosiect ddechrau. Pob dymuniad da i ti oddi wrth y tîm i gyd: paid bod yn ddierth!

+ Cysylltu â ni

Mae'r wybodaeth ddiweddaraf am ddatblygiadau'r prosiect hefyd ar gael drwy Facebook: www.facebook.com/CorCenCC/; Twitter https://twitter.com/corcencc (gallwch ein trydar @CorCenCC). Gallwch hefyd gysylltu â ni drwy anfon neges i gyfeiriad ebost y prosiect: corcencc@caerdydd.ac.uk neu ewch i'n gwefan, sef: www.corcencc.cymru

Mae CorCenCC yn brosiect ymchwil a ariennir gan ESRC/AHRC (Grant Rhif ES/M011348/1). Mae tîm CorCenCC yn cynnwys y Prif Ymchwilydd - Dawn Knight; y Cyd-Ymchwilwyr - Tess Fitzpatrick, Steve Morris, Irena Spasić, Paul Rayson, Enlli Thomas, Mark Stonelake a Jeremy Evas; y Cynorthwywyr Ymchwil - Steven Neale, Jennifer Needs, Mair Rees, Scott Piao a Gareth Watkins; y myfyriwr PhD - Vigneshwaran Muralidaran; Ymgynghorwyr - Kevin Donnelly, Kevin Scannell, Laurence Anthony, Tom Cobb, Michael McCarthy a Margaret Deuchar; Grŵp Ymgynghorol y Prosiect - Colin Williams, Karen Corrigan, Llion Jones, Maggie Tallerman, Mair Parry-Jones, Gwen Awbery, Emyr Davies (CBAC-WJEC), Gareth Morlais (Llywodraeth Cymru), Owain Roberts (Llyfrgell Genedlaethol Cymru), Aran Jones (Saysomethingin.com) ac Andrew Hawke (Geiriadur Prifysgol Cymru). Os oes gennych unrhyw sylwadau neu gwestiynau am gynnwys y cylchlythyr hwn, cysylltwch â Dr Dawn Knight: KnightD5@caerdydd.ac.uk