Mae'r Gymraeg yn iaith adnodd-isel ar y rhyngrwyd, sy'n golygu bod llai o adnoddau a data Cymraeg ar gael i'w defnyddio wrth hyfforddi systemau deallusrwydd artiffisial. Yn syml, nid yw offer AI yn “cyfieithu” y Gymraeg i’r Saesneg, ond yn hytrach yn trosi iaith yn ddata. Mae geiriau a brawddegau yn cael eu cynrychioli fel patrymau mathemategol y gall y model eu prosesu. Fodd bynnag, gan fod llawer mwy o ddata Saesneg ar gael i hyfforddi modelau AI, mae’r systemau hyn yn aml yn meddu ar ddealltwriaeth ddyfnach o’r Saesneg nag o ieithoedd lleiafrifol megis y Gymraeg. Mae’r offeryn AI yn defnyddio algorithmau i ddod o hyd i'r ymateb mwyaf ‘tebyg’ yn seiliedig ar yr wybodaeth a gofnodir. Oherwydd mynychtra cynnwys Saesneg, mae’n anochel y gall gwybodaeth Saesneg gael ei chyrchu yn gyntaf.

O ganlyniad i hyn, gall fod gan yr wybodaeth ragfarn Eingl-ganolog, ac efallai na fydd wedi cyrchu gwybodaeth yn yr iaith Gymraeg.

Lle mae bylchau yn y chwiliad cychwynnol hwn, mae’n debygol y bydd yr offer AI yn creu 'gwybodaeth' credadwy i lenwi bylchau – gwybodaeth sy'n ymddangos yn gredadwy, ond sy'n anghywir neu'n ffug.

Beth allwn ni ei wneud?

Gwirio
Fel gyda phob allbwn a gynhyrchir gan AI, dylid gwirio popeth. Gan fod y Gymraeg yn iaith leifrifol, mae'r cam hwn yn bwysicach fyth.

Peirianneg Promptio. Wrth gwrs, byddwch am ddefnyddio'r amrywiaeth ehangaf o ffynonellau da wrth ymchwilio, ond i sicrhau bod gennych ffynonellau Cymraeg, gallwch fod yn fwy penodol gyda'ch awgrymiadau (lleoleiddio ysgogiadau) – peidiwch â phoeni os ydyn nhw'n hir!

(*Dylid cadw promptiau mor gryno â phosibl. Nid oes angen cwrteisi na rhagymadroddion, gan fod hyn yn cynyddu costau ac effaith amgylcheddol heb fudd ymarferol.)*

1. **<u>Promptio yn Gymraeg!</u>**

Mae tystiolaeth i ddangos y gall ysgrifennu promptiau mewn ieithoedd lleiafrifol arwain at chwiliadau neu ymatebion llai Eingl-ganolog. O bosib, gall hyn fod yn effeithiol pan fyddwch yn cyfuno hyn gyda "chwiliad dwfn".

2. **<u>Gofynnwch i AI chwilio ffynonellau Cymraeg – byddwch yn benodol:</u>**

e.e. *Ateba gan ddefnyddio DIM OND FFYNONELLAU CYMRAEG.*
Rhaid cynhyrchu ffynonellau yng Nghymru (e.e. Llywodraeth Cymru, S4C, prifysgolion yng Nghymru). Rho ddolen uniongyrchol i bob ffynhonnell.

CEFNOGWYD GAN
COLEG CYMRAEG CENEDLAETHOL
Prifysgol Cymru Y Drindod Dewi Sant
University of Wales Trinity Saint David
Peniarth .cymru
RHAGOR iaith
CANOLFAN GWASANAETHAU CYMRAEG
WELSH LANGUAGE SERVICES CENTRE

3. Enwi sefydliadau dibynadwy Cymru yn benodol

Mae offer AI yn dilyn angorau diriaethol yn well na cheisiadau haniaethol.

e.e. *Blaenoriaethu ffynonellau Llywodraeth Cymru*

Sefydliadau sy'n gweithio'n dda

- Llywodraeth Cymru / Welsh Government
- Senedd Cymru
- S4C / BBC Cymru
- Llyfrgell Genedlaethol Cymru
- Prifysgolion: Y Drindod Dewi Sant, Bangor, Aberystwyth, Caerdydd, Abertawe
- Coleg Cymraeg Cenedlaethol
- Estyn
- Mentrau Iaith Cymru

3. Gorfodi hidlo iaith yn benodol

Rhagosodiad (default) offer AI yw'r Saesneg oni bai eich bod chi'n gosod cyfyngiadau. Cofiwch, nid person yw AI! Mae angen i chi roi cyfarwyddiadau clir, penodol.

Cyfyngiad llym

e.e. *Rhaid i bob dyfyniad fod yn Gymraeg.*
Peidiwch â chyfieithu ffynonellau Saesneg i'r Gymraeg.

Cyfyngiad meddal (os yw ffynonellau'n brin)

e.e. Os nad oes ffynhonnell Gymraeg yn bodoli, nodwch hyn yn benodol cyn defnyddio unrhyw ffynhonnell Saesneg.

4. Gofyn am ddyfyniadau dilys, gwir

Mae hyn yn wir am bob defnydd o AI - waeth beth fo'r iaith! Heb hyn, gall y model rhithweld (hallucinate) ffynonellau "Cymraeg" neu greu “gwybodaeth” gan ddefnyddio patrymau ac enghreifftiau.

Gallwch ofyn i AI ddyfynnu pob allbwn gyda:

enw'r sefydliad
Teitl Cymraeg y ddogfen / gwefan / ffynhonell
URL uniongyrchol

e.e. *Rhowch ddyfyniadau troednodiadau gyda dolenni uniongyrchol i dudalennau Cymraeg.*
Peidiwch â dyfeisio na brasamcanu ffynonellau.
Defnyddiwch ffynonellau gwirioneddol yn unig.

Cofiwch! Dylech bob tro wirio’r ffynonellau a gynigir gan yr offer.

5. Eithrio ffynonellau diangen

Gall cyfyngiadau eithrio hefyd fod yn effeithiol iawn.

e.e. *Eithriwch gyfnodolion academaidd Saesneg.*
Peidiwch defnyddio...

6. Gofynnwch i'r AI hunanwirio

Mae hyn yn lleihau 'drifft' model. Yn syml ar gyfer ein cyd-destun ni, drifft model AI yw'r duedd i atebion lithro i ragdybiaethau Saesneg ac Eingl-Ganolog (mae'n drifftio'n ôl i'r canol) oni bai bod y cais wedi'i angori'n gadarn yn y Gymraeg.

e.e. Cyn ateb, gwiriwch yn fewnol:

1. Mae'r holl ffynonellau yn Gymraeg

2. Mae'r holl ffynonellau yn tarddu o Gymru

3. Mae'r holl ddolenni yn hygyrch, ac yn ddilys

Cofiwch, gall systemau AI arddangos drifft yn ystod deialog, gan oresgyn yn raddol ofynion ieithyddol a bennwyd gan y defnyddiwr a dychwelyd at osodiadau diofyn (yn aml y Saesneg). Dylech fod yn wyliadwrus ac ail-bromptio os oes angen.

7. Cael prompt craidd wrth law

Gallwch ailddefnyddio hwn bron bob amser:

Rhowch ateb gan ddefnyddio DIM ond ffynonellau Cymraeg.

Gofynion:

- Rhaid i ffynonellau ddod o Gymru

- Blaenoriaethu Llywodraeth Cymru, Senedd Cymru, S4C, BBC Cymru, Estyn, neu brifysgolion Cymru

- Dyfynnu pob honiad ffeithiol gydag URL Cymraeg uniongyrchol

- Peidiwch â chyfieithu ffynonellau Saesneg i'r Gymraeg (bydd hyn yn osgoi rhithwelediadau terminolegol).

- Os nad oes ffynonellau Cymraeg addas yn bodoli, nodwch hyn yn glir

Pwnc: [eich pwnc yma]

Mae'r promptiau yn y canllaw hwn wedi'u cynllunio i'ch helpu i gael mwy o gynnwys a ffynonellau Cymraeg. Fodd bynnag, oherwydd natur AI, gall yr allbwn fod yn anghywir neu'n ddychmygol ar adegau. Dylid felly ystyried AI fel cymorth i'ch dysgu—nid fel awdurdod. Gwiriwch bob ffaith bob amser cyn ei defnyddio neu ei chyflwyno.

CEFNOGWYD GAN
COLEG CYMRAEG CENEDLAETHOL

Prifysgol Cymru Y Drindod Dewi Sant
University of Wales Trinity Saint David

Peniarth.cymru

RHAGOR iaith

CANOLFAN GWASANAETHAU CYMRAEG
WELSH LANGUAGE SERVICES CENTRE