

Tuag at Gynorthwydd Personol Deallus Cymraeg

**Astudiaeth Fer o APIs ar gyfer Gorchmynion
Llafar, Systemau Cwestiwn ac Ateb a Thestun
a Lleferydd ar gyfer Llywodraeth Cymru**

Stefano Ghazzali

Dewi Bryn Jones

Delyth Prys

Hydref 2015

Uned Technolegau Iaith, Canolfan Bedwyr, Prifysgol Bangor

<http://techiaith.bangor.ac.uk>



Llywodraeth Cymru
Welsh Government



Cynnwys

1 Crynodeb

2 Cefndir

3 Methodoleg

4 Canlyniadau

4.1 Pensaernïaeth Cynorthwydd Digidol Generig

4.2 Microsoft Cortana

4.2.1 Project Oxford

4.3 Google Voice Actions

4.4 Siri Apple

4.5 Alexa Amazon Echo

4.6 Casgliadau Pensaernïaeth Gaeedig

4.7 Dewisiadau Cod Agored

4.7.1 Project Jasper

4.7.2 SIRIUS

4.7.3 Wit.ai

4.7.4 Mycroft.ai

5. Casgliadau

6. Cyfeiriadau

1 Crynodeb

Mae'n gynyddol bosibl i ddefnyddwyr siarad mewn Saesneg naturiol gyda'u dyfeisiau a'u cyfrifiaduron er mwyn cwblhau rhai tasgau, cael cymorth neu ofyn am wybodaeth.

Mae'r astudiaeth hon yn adolygu'r cynorthwyyr personol deallus sydd ar gael heddiw, eu pensaernïaeth fewnol o ran cydrannau technoleg iaith a'r cyfleoedd maent yn rhoi i gefnogi defnyddwyr sydd â dewis iaith wahanol, sef, yn ein hachos ni, y Gymraeg.

Daw'r astudiaeth i'r casgliad fod pensaernïaethau cynorthwyyr personol deallus poblogaidd yn rhy gaeedig at ein dibenion ni, a bod eu citiau datblygu a'u APIs yn rhy gyfyngedig i'w gwneud hi'n ddichonadwy i ymchwilwyr a datblygwyr eu mabwysiadu i gefnogi iaith naturiol Gymraeg. O'r herwydd mae'r astudiaeth hon yn argymhell dulliau y gall y Project Seilwaith Cyfathrebu Cymraeg fwrw ymlaen i wireddu Cynorthwyydd personol deallus gyda llwyfannau a pensaernïaethau agored.

2 Cefndir

Mae'n gynyddol bosibl i chi siarad gyda dyfeisiau fel eich ffôn neu gyfrifiadur er mwyn gorchymyn a rheoli rhaglenni a dyfeisiau, a hefyd dderbyn atebion deallus a pherthnasol i gwestiynau a ofynnwyd mewn iaith naturiol.

Mae'r galluoedd hyn yn bosibl o ganlyniad i gynnydd diweddar mewn adnabod lleferydd, cyfieithu peirianyddol a phrosesu a deall iaith naturiol. Hwyl felly yw'r prif alluogwyr ar gyfer newid a fydd yn tarfu ar y drefn bresennol ac yn achosi shifft sylfaenol yn y ffordd y bydd defnyddwyr yn ymgysylltu â'u dyfeisiau a'r ffordd y byddant yn defnyddio technoleg yn ehangach.

Os edrychir ar hyn yn ei gyd-destun hanesyddol ehangach, hwn yw'r cam nesaf naturiol yn natblygiad y rhyngweithio rhwng pobl a chyfrifiaduron; o'r bysellfwrdd i'r llygoden, i dechnoleg cyffwrdd, i lais ac i iaith.

Mae pedwar prif lwyfan masnachol yn gyrru'r newid hwn, sef Siri, OK Google, Microsoft Cortana and Amazon Alexa, yn ogystal â rhai llwyfannau cod agored llai adnabyddus. Hyd yn hyn mae'r rhain yn darparu eu galluoedd pwerus yn Saesneg a rhai ieithoedd mawr eraill, a phrin yw'r dystiolaeth eu bod yn debygol o ymestyn eu dewis o ieithoedd i gynnwys y 'gynffon hir' o ieithoedd llai, gan gynnwys y Gymraeg, yn y dyfodol agos.

Noddwyd y project hwn felly gan Lywodraeth Cymru drwy ei Chronfa Technoleg Gymraeg a'r Cyfryngau Digidol ac S4C er mwyn sicrhau nad yw defnyddwyr sydd â'r Gymraeg yn ddewis iaith iddynt yn cael eu gadael ar ôl mewn datblygiadau o'r fath. Bydd yn gosod sylfeini amrediad o dechnolegau Cymraeg i'w defnyddio yn yr amgylcheddau hyn, gan gynnwys gwella'r gwaith a wnaeth hyd yn hyn ar adnabod lleferydd Cymraeg yn ogystal â chyfieithu peirianyddol er mwyn cael y budd mwyaf allan o'r galluoedd a ddarperir drwy dechnolegau wedi'u seilio ar y Saesneg. Bydd y

project yn symbylu datblygiad meddalwedd a gwasanaethau Cymraeg newydd a allai gyfrannu at brif-ffrydio'r Gymraeg yn y cam nesaf o ryngweithio rhwng pobl a chyfrifiaduron.

3 Methodoleg

Bu i ni edrych i ddechrau ar y cynorthwyr deallus sy'n cael eu cynnig gan y pedwar cwmni mwyaf gweithgar yn y maes, sef Google, Microsoft, Apple ac Amazon. Ymchwiliwyd i ddogfennaeth gysylltiedig a oedd ar gael ar gyfer datblygwyr ochr yn ochr ag unrhyw APIs ac SDKs er mwyn asesu unrhyw botensial i'w haddasu ar gyfer defnyddwyr Cymraeg eu hiaith.

O chwilio am rai allweddeiriau sylfaenol ar y we daethpwyd o hyd i nifer o brojectau cod agored perthnasol yn ogystal â rhai gwasanaethau API ar-lein llai adnabyddus. Gwerthuswyd y rhain drwy edrych ar eu dogfennaeth a'u cod ffynhonnell i weld a ellid eu haddasu ar gyfer defnyddwyr Cymraeg.

Lluniwyd pensaernïaeth generig (gw. adran 4.1) a oedd yn gyffredin i bob cynorthwyydd deallus ar sail hyn, a'i ddefnyddio i werthuso pob llwyfan yn llawn o ran ei ddichonoldeb a'r cyfleoedd i'w fabwysiadu ar gyfer y Gymraeg.

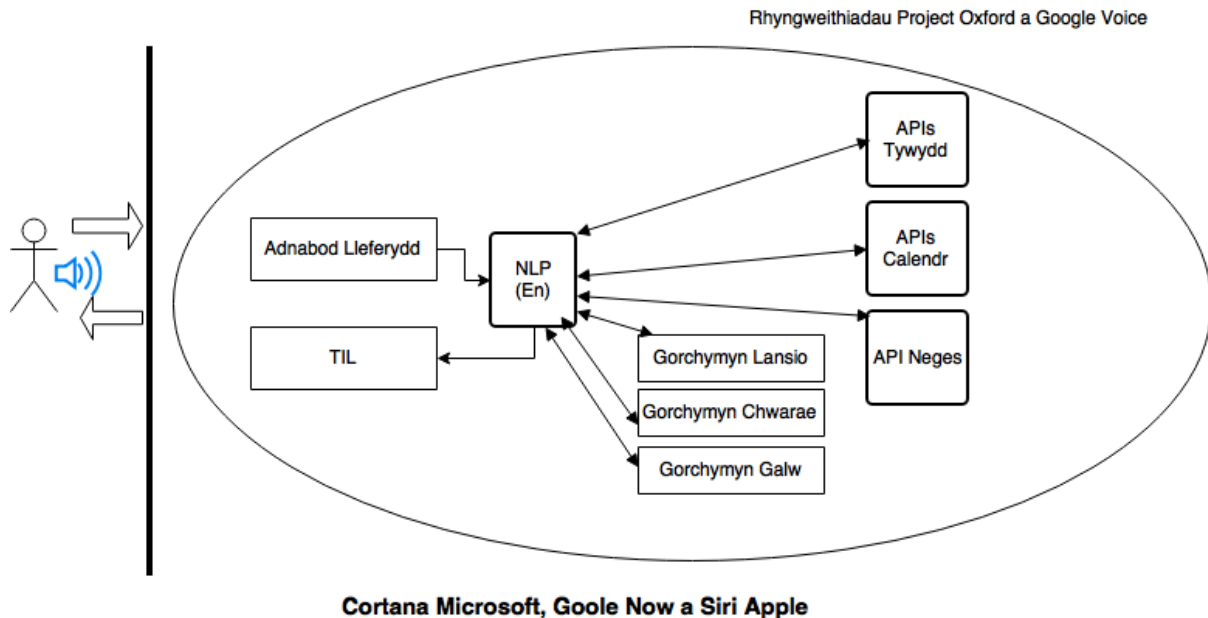
Esgorodd hyn ar strategaeth i gyflawni amcanion y Project Cyfathrebu Cymraeg ar ôl cwblhau'r ymarferiad gwerthuso.

Er mwyn cadw'r astudiaeth hon yn syml, ni roddwyd pwyslais ar yr iaith y cyflwynid yr ymatebion a'r canlyniadau ynddi.

4 Canlyniadau

4.1 Pensaernïaeth Cynorthwydd Personol Digidol Generig

Mae'r ffigur canlynol yn darlunio'r bensaernïaeth sylfaenol y mae pob cynorthwydd deallus yn ei rhannu.



Mae defnyddiwr yn siarad Saesneg naturiol gyda'r ddyfais neu'r cyfrifiadur.

1. Mae peiriant adnabod lleferydd Saesneg yn cipio'r awdio ac yn trosi dymuniad llafar y defnyddiwr i destun sy'n cael ei drosglwyddo i gydran prosesu iaith naturiol (NLP) Saesneg.
2. Gall y gydran NLP brosesu'r testun a deall beth yw bwriad neu ddymuniad y defnyddiwr ac felly gall adnabod pa gydran neu API fydd yn cyflawni'r cais (efallai gyda chymorth pob API neu gydran).
3. Bydd y gydran NLP yn creu neges gyda'r gwahanol ymresymiau i gyfathrebu gyda modiwl yn ôl manyleb ei API.
4. Mewn achosion lle mae angen ateb, bydd yr NLP yn derbyn ymateb oddi wrth yr API sy'n cyflawni'r weithred ac yn cynhyrchu brawddeg sy'n cynnwys y canlyniadau.
5. Caiff y frawddeg mewn iaith naturiol ei throsglwyddo i beiriant testun i leferydd Saesneg i'w llefaru yn ôl wrth y defnyddiwr.

Er enghraifft, gall y defnyddiwr ofyn "Do I have any meetings today?"

1. Mae'r peiriant adnabod lleferydd yn trosi'r awdio i'r testun 'DO I HAVE ANY MEETINGS TODAY' ac yn ei anfon at y gydran NLP.
2. Mae'r gydran NLP yn prosesu'r testun ac yn ei adnabod fel cwestiwn 'DO I HAVE', yn adnabod 'MEETINGS' a 'TODAY' fel allweddeiriau sy'n gysylltiedig ag amser a chalendr. Mae'r NLP yn holi'r API calendr i gael

cadarnhad, ac mae'r ddau yn cytuno.

3. Mae'r NLP yn llunio neges yn ôl manyleb yr API Calendr ac yn ei anfon e.e.

```
getEvents(type=meeting, date=today);
```

4. Mae'r NLP yn derbyn rhestr mewn ymateb, efallai ar ffurf megis JSON:

```
{"success": true, "events": [ {"date": "2015-09-25",  
"time": "10:30am", "name": "Meeting to discuss a new  
Welsh language project", "location": "Bangor  
University, Ogwen Building, Room 234"}, {"date":  
"2015-09-25", "time": "12:30pm", "name": "Lunch with  
Delyth", "location": "Terrace Restaurant, Bangor  
University"} ]}
```

5. ac yn cynhyrchu'r frawddeg isod ohono:

“Yes you do. At 10:30 this morning you have a meeting to discuss a new Welsh language project in Bangor University, Ogwen Building, Room 234. Then at 12:30 you have lunch with Delyth in the Terrace Restaurant, Bangor University”

6. Mae'r peiriant testun i leferydd yn derbyn y canlyniad ar ffurf brawddeg, ac yn lleisio:

“Yes you do. At half past ten this morning you have a meeting to discuss a new Welsh language project in Bangor University, Ogwen Building, Room two hundred and thirty four. Then at half past twelve you have lunch with Delyth in the Terrace Restaurant, Bangor University”

4.2 Microsoft Cortana

Wrth ryddhau Windows 10, mae Microsoft wedi rhyddhau ei ap cynorthwydd personol gyda'r gallu i lefaru dan yr enw Cortana. Nid yw Cortana'n gyfyngedig i gyfrifiaduron a dyfeisiau Windows 10, ac mae ar gael hefyd ar gyfer Android ac iOS

Dywed tudalen gartref datblygwyr Cortana (<http://dev.windows.com/en-us/cortana>):

“While real-world personal assistants do a lot by themselves, they often rely on a network of experts to help with specific tasks.

In much the same way, Cortana knows her users and what they're trying to do, and can use the apps on a user's device as her own network of experts. Make your app part of that network.”

Mae Microsoft yn darparu ffordd i ddatblygwyr integreiddio eu hapiau gyda Cortana fel y gallant ddarparu rhyngwyneb lleferydd deallus i'w hapiau. Mae amrediad galluoedd Cortana hefyd yn cael ei ymestyn gydag apiau integredig sy'n debyg yn

ddamcaniaethol i'r modiwl APIs a welir yn Ffigur 1.

Mae Microsoft yn mynd yn bellach wrth ddarparu gwasanaethau testun i leferydd a 'Diffiniadau Gorchmynion Llais' ('*Voice Command Definitions*' neu VCDs), sef ffeiliau diffinio XML syml i ymestyn Cortana y gall datblygwyr eu defnyddio i adeiladu eu rhyngweithiadau lleferydd iaith naturiol eu hunain sy'n dechrau gyda Cortana.

Mae pensaernïaeth Cortana yn amgáu'r adnabod lleferydd a'r prosesu a'r deall iaith naturiol o fewn un uwch-gydran. Nid yw'r bensaernïaeth yn gwahanu'r ddwy swyddogaeth a rhoi mynediad API ar wahân i'r ddwy gydran, yn arbennig ei chydrannau prosesu iaith naturiol.

Byddai cyfleuster o'r fath yn galluogi cynorthwydd personol deallus Cymraeg i integreiddio Cortana y tu ôl i beiriant adnabod lleferydd Cymraeg a chydrannau cyfieithu peirianyddol.

Fodd bynnag, rhaid i bob rhyngweithiad lleferydd drwy Cortana gychwyn gydag awdio yn yr ieithoedd y dywedir fod Cortana yn eu cynnal. Nid oes modd ystyried Cortana Microsoft ymhellach yn y project hwn.

4.2.1 Project Oxford Microsoft

Mae gan Microsoft ystod o APIs ac SDKs dechnolegau iaith yn y cwmwl wedi'u lleetya gyda ProjectOxford.ai sy'n gwasanaethu fel technolegau craidd Cortana. Maent yn darparu ffordd i ddatblygwyr integreiddio technolegau Microsoft mewn sefyllfaoedd a datrysiadau uwch na'r rhai a ganiateir drwy Cortana yn unig.

Yn benodol:

- APIs Lleferydd
 - Adnabod Lleferydd – trosi awdio llafar i destun (fel canlyniadau adnabod llawn neu rannol) mewn amser real.
 - Adnabod Bwriad Lleferydd – yn ychwanegol at ddychwelyd testun o awdio, mae'r API yn dychwelyd gwybodaeth strwythuredig ar unrhyw fwriad. Gall eich ap barsio a phenderfynu ar unrhyw weithredu pellach.
 - Testun i Lleferydd – peiriant testun i leferydd yn y cwmwl.
- Gwasanaeth Deallus Adnabod Iaith (*Language Understanding Intelligent Service* neu LUIS) – sy'n gadael i chi ddefnyddio adnabod bwriad heb adnabod lleferydd. Yn ychwanegol mae'n galluogi datblygwyr i greu bwriadau newydd a modelu sut mae rhai traethiadau iaith naturiol yn mapio ar i ddata a pharsio bwriad strwythuredig yn apiau datblygwyr.

Nid yw LUIS yn cynnal ffurfio modelau bwriad a fynegwyd yn Gymraeg. Fodd bynnag, gellid integreiddio'r gwasanaeth i brosesu canlyniadau cyfieithu peirianyddol Cymraeg i Saesneg.

Ar hyn o bryd mae'r holl APIs ar gael ar ffurf beta preifat a hynny drwy wahoddiad yn unig. Nid ydynt ar gael felly i'w hystyried ymhellach o fewn y project hwn.

4.3 Google Voice Actions

Mae Google Voice Actions yn caniatáu i ddatblygwyr i ychwanegu galluoedd gorchmynion llais i'w hapiâu ar ffonau, tabledi, setiau teledu a dyfeisiau Wear Android.

Caiff gorchmynion llais eu llefaru i Google Search a'u cyfeirio at apiau yn ôl y brawddegau a'r ymadroddion sydd wedi'u ffurfweddu i danio gweithredoedd penodol. Er enghraifft, byddai "Ok Google, turn on the lights on MyApp" yn peri i'r ap MyApp agor gyda'i API a rhoi cyfarwyddiadau i osod y goleuadau i fod ynghyn. [3]

Gall Voice Search Google gynnal 50 iaith yr un pryd, ond nid y Gymraeg. Mae gwasanaeth adnabod lleferydd ar gyfer Google Voice Actions, a'i wasanaeth "OK, Google" hefyd yn ehangu ei gefnogaeth i ieithoedd. Fodd bynnag, mae cefnogaeth i'r Gymraeg, heb gefnogaeth sylfaenol yn Voice Search, i weld yn bell i fwrdd.

Mae arlwy Voice Actions Google i ddatblygwyr felly yn debyg o ran ei adeiladwaith a'i rychwant i eiddo Cortana Microsoft. Felly mae'n syrthio'n fyr o ran hyblygrwydd a bod yn agored, ac ni ellir ei ystyried ymhellach i wireddu cynorthwydd personol deallus Cymraeg.

4.4 Siri Apple

Cynorthwydd personol Apple yw Siri, ac fe'i ceir yn ei ddyfeisiau sy'n seiliedig ar iOS ers nifer o flynyddoedd, ac mae wedi dod yn rhan annatod o ddefnydd dyddiol pobl o'u iPhones a'u iPads.

O'i gyferbynnu â chwmnïau eraill sy'n cynnig llwyfannau cynorthwywyr personol, nid yw Apple wedi rhannu gwybodaeth am fynediad agored nac yn rhannu technoleg Siri gyda datblygwyr, er rhwystredigaeth a dicter datblygwyr. Mae Siri yn hollol gaeedig, ond yn rhannu'r wybodaeth fod Siri yn defnyddio casgliad o tua 35 API i gyflawni'r tasgau gofynnol.

Gyda rhyddhau iOS 9 ac El Capitan fodd bynnag, mae Apple yn caniatáu i apiau iOS gysylltu'n ddwfn gyda Search a Siri, mewn ffordd a rychwant digon tebyg i eiddo Voice Actions Google. Mae API Search newydd iOS 9 yn galluogi datblygwyr i greu allweddeiriau y bydd adnabod lleferydd Siri ei hun yn eu hadnabod i danio unrhyw weithgaredd cysylltu dwfn.

Nid yw'r datblygiadau hyn o fewn iOS Siri yn ddigon agored a hyblyg o hyd i gynorthwyo datblygu cynorthwydd personol Cymraeg, ac felly ni ellir ystyried Siri ymhellach yn y project hwn.

4.5 Alexa Amazon Echo

Dim ond ers Gorffennaf 2015 mae Amazon Echo wedi bod ar werth yn yr Unol Daleithiau. Mae ymateb cychwynnol defnyddwyr wedi bod yn gadarnhaol iawn am ei alluoedd gyda llawer o arddangosiadau fideo ar gael ar wefannau megis YouTube o

Alexa yn mynd trwy'i bethau ac yn ymateb yn ddefnyddiol i gwestiynau heriol. Hyd yn hyn ni chafwyd unrhyw gyhoeddiad gan Amazon ynghylch pryd fydd Echo ar gael yn Ewrop ac yn y DU.

Alexa yw'r gair sy'n deffro Echo Amazon a'r enw a roddir i'r cynorthwydd llais personol sydd, er ei fod y mynediad iddo drwy Echo, yn cael ei bweru a'i letya yn isadeiledd cwmwl Amazon.

I gyd-fynd gyda rhyddhau'r ddyfais a'r gwasanaethau i ddefnyddwyr, rhyddhaodd Amazon hefyd y Cit Sgiliau (*Alexa Skills Kit* neu ASK) a Gwasanaeth Llais Alexa (*Alexa Voice Service* neu AVS) fel ffordd i ddatblygwyr weithio gydag Alexa.

Mae Gwasanaeth Llais Alexa yn galluogi gwneuthurwyr caledwedd dyfeisiau sy'n medru cysylltu i'r rhyngwyd, ac sydd â microffon a seinydd, i ychwanegu nodwedd cynorthwydd Alexa i'w dyfeisiau. Gellid integreiddio gwasanaeth o'r fath i apiau ar iOS.

Yn y cyfamser mae'r Set Sgiliau Alexa yn galluogi datblygwyr i ehangu galluoedd Alexa drwy ddiffinio bwriadau y gall y gwasanaeth Alexa eu deall o destunau iaith naturiol fel y mae adnabod lleferydd Amazon yn eu hadnabod, ac yna integreiddio eu APIs er mwyn gwasanaethu eu bwriadau.

Hyd yn hyn mae Alexa ar gael yn unig yn Saesneg America ac nid yw'r SDKs a ddisgrifir uchod yn galluogi datblygu cynorthwydd personol Cymraeg ar sail Alexa.

Y mae Amazon fodd bynnag yn darparu ffordd a chronfa (Cronfa Alexa gwerth \$100 miliwn) i ddatblygwyr a chwmnïau arloesol gyfrannu at wella'r technolegau gwaelodol, sef adnabod lleferydd, deall iaith naturiol a thestun i leferydd. Mae'r Uned wedi gwneud ceisiadau cychwynnol i ychwanegu cefnogaeth Gymraeg ac amlieithog i Alexa.

Fel arall, nid ystyrir Alexa Amazon ymhellach yn y project hwn.

4.6 Casgliadau Pensaerniaeth Gaeedig

Rydym wedi dangos fod gan y llwyfannau cynorthwywyr digidol deallus a ddarperir gan y pedwar cwmni mwyaf eu maint a mwyaf gweithgar bensaerniaeth gaeedig. Golyga hyn fod mynediad ar gyfer ymchwilwyr a datblygwyr wedi'i gyfyngu, ac mae'n canolbwyntio yn bennaf ar y rhai a all ymestyn galluoedd y cynorthwydd drwy integreiddio i'w hapiâu a'u APIs eu hun.

Mae addasu er mwyn cynnal mewnbwn prosesu a lleferydd ieithoedd eraill yn amhosibl ym mhob enghraifft.

O ganlyniad, ni fydd modd integreiddio apiau a gwasanaethau Cymraeg er mwyn ymestyn galluoedd cynorthwywyr i ddefnyddwyr sydd angen cymorth penodol yn y Gymraeg.

Fodd bynnag, nid yw hyn yn cau allan y posibilrwydd y gall cwmnïau mawr a gweithgar fel hyn ymestyn eu dewis o ieithoedd i gynnwys ieithoedd llai megis y

Gymraeg drwy sylweddoli'n gynyddol er y gall fod marchnadoedd yn y 'gynffon hir' o ieithoedd llai fod yn fychan ar eu pen eu hyn, eu bod gyda'i gilydd yn gyfran sylweddol o ddefnyddwyr a chwsmeriaid posibl. Gall modiwlareiddio cydrannau a darpariaeth adnoddau anfoddol ar gyfer datblygu'r technolegau gwaelodol mewn storfeydd sydd â mynediad rhydd iddynt megis Porth Technolegau Iaith Cenedlaethol Cymru, fel a ragwelir yn y project presennol, hefyd annog y cwmnïau mwy i'w mabwysiadu.

4.7 Dewisiadau Cod Agored

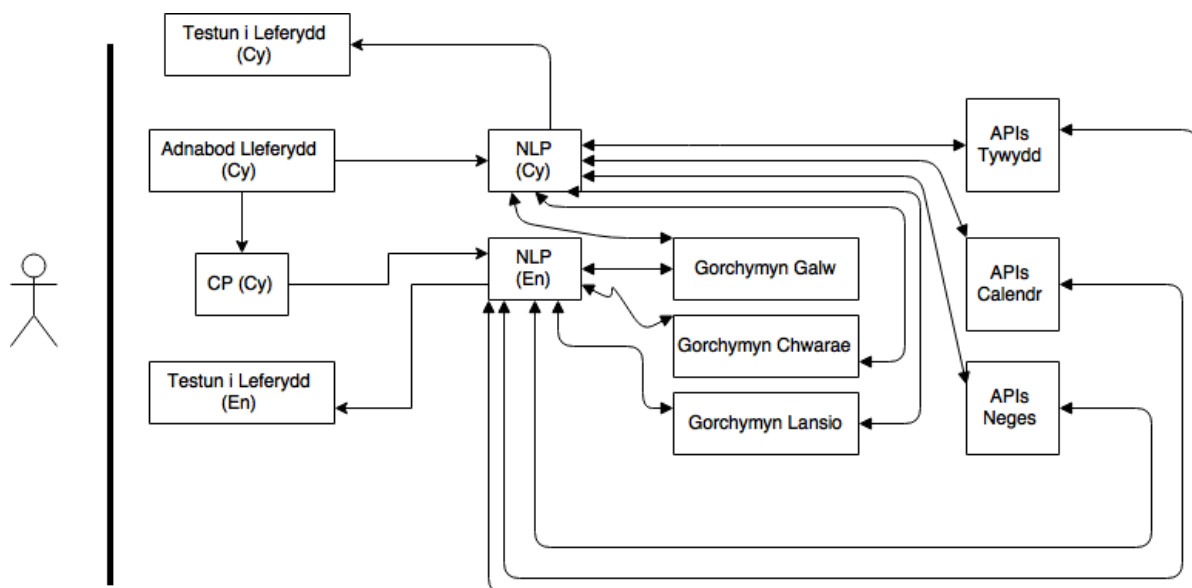
Mae angen pensaernïaethau llawer mwy agored a hyblyg ar gynorthwydd digidol personol Cymraeg lle gellir integreiddio nifer o dechnolegau iaith.

Mae nifer o lwyfannau cynorthwywyr personol deallus yn bodoli sy'n fwy agored o ran pensaernïaeth. Gwerthuswyd APIs a chod ffynhonnell fel y brif ffynhonnell i gael datrasiad.

Ein meini prawf oedd:

- Y gellid integreiddio peiriant adnabod lleferydd Cymraeg
- Y gellid naill ai addasu neu ddisodli'r NLP ar gyfer deall iaith
- Y gellid darparu ymatebion drwy destun i leferydd Cymraeg
- Y gellid dal i gynnwys APIs a modiwlau sy'n cyflawni tasgau ond sy'n seiliedig ar y defnydd o'r Saesneg, gan integreiddio cyfieithu peirianyddol Cymraeg i Saesneg at ddibenion newydd.

Mae'r ffigur dilynol yn darlunio pensaernïaeth bosibl:



4.7.1 Jasper

Mae Project Jasper yn broiect cod agored syml iawn sydd wedi'i ddatblygu'n bennaf i redeg Raspberry Pi wedi'i gysylltu wrth y rhyngwyd gyda dim ond microffon ac uchelseinydd ynghlwm wrtho [8]. Defnyddir y bensaernïaeth gyffredinol a welir yn y ffigur yn adran 4.1 i'w wireddu.

Mae'n hawdd integreiddio cydrannau sy'n cynnig swyddogaeth benodol. Felly mae'n galluogi'r datblygwr i ddewis pa adnabod lleferydd gyda chod integreiddio sydd ar gael yn barod ar gyfer adnabod lleferydd ar ddyfais fel PocketSphinx a Julius/HTK neu yn y cwmwl fel y rhai y mae Google yn eu darparu. Gellir gwneud newidiadau syml i'r cod i gynnwys cydrannau eraill megis cyfieithu peiranyddol.

Mae cydran 'ymennydd' yn darparu dull i ymestyn galluedd Jasper yn hawdd lle gall, ynghyd â modiwlau integredig neu APIs, ddehongli a deall bwriadau.

Gellir hefyd ychwanegu cydrannau testun i leferydd er bod Jasper hefyd yn integreiddio gyda Pheiriant Synthesis Lleferydd Festival – peiriant y mae gan yr Uned eisoes brofiad ohono a llais Cymraeg yn seiliedig arno.

Mae Jasper yn atyniadol am fod iddo bensaernïaeth syml iawn ar gyfer gwaith datblygu a phrawf o'r cysyniad sy'n gysylltiedig â datblygu'r cydrannau technolegau iaith sydd dan y cyfan. Mae hefyd yn atyniadol iawn i'w ddefnyddio mewn cyd-destun addysgol a chlybiau codio a allai ddatblygu ac ymestyn galluedd.

4.7.2 SIRIUS

Cynorthwydd personol deallus (*intelligent personal assistant* neu IPA) yw Sirius sydd yn god agored pen-i-ben ac yn sefyll ar ei ben ei hun, yn seiliedig ar dechnoleg lleferydd a golwg. Mae'r project yma wedi'i seilio yn y Clarity Lab ym Mhrifysgol Michigan. Nid yw'n gysylltiedig unrhyw ffordd â Siri Apple.

Mae Sirius yn derbyn ymholiadau ar ffurf lleferydd neu ddelweddau ac yn dychwelyd canlyniadau ar ffurf iaith naturiol. Mae Sirius yn gweithredu swyddogaethau craidd IPA gan gynnwys adnabod lleferydd, cydweddu delweddau, prosesu iaith naturiol a system cwestiwn ac ateb. [9]

Mae Sirius wedi'i ddylunio i'w osod ar beiriannau neu weinyddion mwy pwerus gyda rhaglenni client enghreifftiol ar gael er mwyn cael mynediad dros y we. Mae gan Sirius eisoes god ar gyfer integreiddio adnabod lleferydd Sphinx CMU.

Gan fod y bensaernïaeth yn ddigon modiwlaid ac agored, mae'n werth defnyddio Sirius fel datrysiad a allai alluogi IPAs Cymraeg yn y cwmwl sy'n hygyrch o unrhyw OS, dyfais a/neu dechnoleg gwisg.

4.7.3 Wit.ai

Cwmni Facebook yw Wit.ai sydd â'r gwaith o adeiladu llwyfannau AI sy'n cynorthwyo datblygwyr i greu apiau sy'n deall iaith ddynol.

Fodd bynnag mae dull Wit.ai yn wahanol i fentrau masnachol eraill gan fod ei wasanaethau API yn fwy modiwlaid. Mae Wit.ai yn honni eu bod yn ymgysylltu gyda chymunedau drwy rannu popeth y mae'n dysgu gyda datblygwyr.

Mae APIs Wit.ai yn y cwmwl yn galluogi datblygwyr i anfon naill ai awdio neu destun cais defnyddiwr ac yn dychwelyd data o'i gydrannau iaith naturiol y mae modd gweithredu arno ac y gellir yn ei dro ei ddefnyddio i gyflawni'r dasg.

Mae Wit.ai o ddiddordeb felly gan fod ei API yn medru derbyn cyfieithu peirianyddol Cymraeg i Saesneg o gais a gafodd ei leisio yn wreiddiol yn Gymraeg.

4.7.4 Mycroft.ai

Project a chynnyrch cod agored yw Mycroft.ai ond y mae'n dal yn fenter fasnachol. Mae'n debyg i Jasper o ran ei fod wedi'i sylfaenu ar y Raspberry Pi ond mae ganddo gasyn gyda goleuadau LED syml wedi'u trefnu i ddynwared wyneb. Ei uchelgais yw 'A.I. i bawb' ac felly mae'n cynnwys prosesu iaith naturiol mwy cymhleth. Yn fyr, mae'n system fwy agored ac estynadwy nag Echo Amazon. Mae Mycroft yn cefnogi nifer o brotocolau y Rhynggrwyd Pethau gan gynnwys NEST, WEMO, Hue ac Iris.

Hyd yn hyn, oherwydd ymgyrchoedd sy'n mynd ymlaen ar hyn o bryd ar Indiegogo, mae'r cynnyrch Mycroft.ai ar gael yn unig fel rhag-archeb ac mae'r cod ffynhonnell wedi'i amserlennu i'w ryddhau wedi Ebrill y 1af 2016. Yn anffodus ni fydd hynny o ddefnydd i'r project presennol.

5. Casgliadau

Dim ond y dewisiadau agored all ddarparu'r hyn sydd ei angen i greu cynorthwydd personol deallus Cymraeg.

Mae'r dewisiadau agored a werthuswyd gennym ni yn amrywio o ran eu soffistigeiddrwydd a'u cymhlethdod ac yn eu cefnogaeth i'r hyn y gallant ei gefnogi.

Mae Jasper yn syml iawn i'w ddefnyddio o fewn Raspberry Pis lle mae Wit.ai a Sirius yn perthyn i'r cwmwl sy'n galluogi llawer o ddefnyddwyr i'w defnyddio o ffynonellau lluosog.

Rydym yn ffodus felly fod y dewisiadau agored yn abl i gynnal strategaeth datblygu cynyddol, lle gellid defnyddio Jasper i ddatblygu system gyda galluoedd un neu ddau barth cyfyngedig megis gofyn yr amser yn Gymraeg, gan ddefnyddio Sirius ar gyfer gofynion mwy cymhleth sy'n delio â cheisiadau cwestiwn ac ateb penagored, a hefyd mynediad drwy apiau a/neu dechnoleg wisgadwy.

Mae dull cam wrth gam fel hyn yn caniatáu i rai o adnoddau eraill yr Uned gael eu cyflwyno i'r broses ddatblygu mewn ffordd reoledig. Er enghraifft, mae gan yr Uned eisoes lais testun i leferydd Cymraeg elfennol, sydd ar gael ar gyfer cynhyrchu lleferydd nail ai o gyfrifiadur (megis y Raspberry Pi) neu drwy gwmwl o wasanaeth APIs Cwmwl yr Uned.

Byddwn hefyd yn medru datblygu adnabod lleferydd i gefnogi nifer cynyddol o geisiadau defnyddwyr posibl, hyd nes, gobeithio, y bydd peiriant adnabod lleferydd di-dor yn cael ei wireddu i dderbyn unrhyw gwestiynau cyffredinol. Eisoes mae'r UTI yn meddu ar nifer o adnoddau perthnasol, y mae wedi'u darparu i eraill drwy'r Porth Technolegau Iaith Cenedlaethol, sy'n 'flociau adeiladu' hanfodol ar gyfer datblygu cynorthwydd personol deallus ar gyfer y Gymraeg. Defnyddir y rhain, a rhai newydd a grëwyd yn benodol yn ystod oes y project hwn, i adeiladu prototeipiau a modelau prawf o gysyniad gan ddefnyddio pensaernïaeth agored. Gallant hefyd gael eu defnyddio gan eraill, gan gynnwys cwmnïau sy'n cadw eu cod eu hunain yn gaeedig, i ymestyn eu cynnyrch i gynnwys y Gymraeg.

Mae'r model agored hwn felly yn cynnig ffordd ymlaen nid yn unig i'r Gymraeg, ond hefydd i gymunedau iaith eraill llai o faint, i greu ecosystem gynhwysol, lle bydd budd cymdeithasol ac economaidd ar gael o rannu adnoddau sylfaenol, lle bydd dyfeisiau a yrrir gan lais yn galluogi siaradwyr pob iaith i ymgysylltu gyda'r genhedlaeth newydd o gyfryngau a chynorthwywyr digidol.

6. Cyfeiriadau

1. Nikola Metulev (2015) *Meet the Speech Platform in Windows 10*, Available at: <http://metulev.com/meet-the-speech-platform-in-windows-10/> (Accessed: 21 September 2015).
2. Microsoft Project Oxford (2015) *Microsoft Project Oxford*, Available at: <https://www.projectoxford.ai/luis> (Accessed: 21 September 2015).
3. Laurence Moroney (2015) *Coffee with a Googler: Learn about Google Voice Actions*, Available at: <http://googledevelopers.blogspot.co.uk/2015/09/learn-about-google-voice-actions.html> (Accessed: 21 September 2015).
4. Google (2015) *Google Voice Interactions - Overview*, Available at: <https://developers.google.com/voice-actions/interaction/> (Accessed: 21 September 2015).
5. ISpeech (2015) *ISpeech API Developer Guide*, Available at: <http://www.ispeech.org/api> (Accessed: 21 September 2015).
6. Politepix UG (2015) *Open Ears Website*, Available at: <http://www.politepix.com/openears/> (Accessed: 21 September 2015).
7. Expect Labs (2015) *Getting Started with the iOS SDK*, Available at: <https://expectlabs.com/docs/sdks/ios/gettingStarted> (Accessed: 21 September 2015).
8. Jasper Project (2015) *Jasper Project Documentation*, Available at: <http://jasperproject.github.io/documentation/> (Accessed: 21 September 2015).

9. Johann Hauswald, Michael A. Laurenzano, Yunqi Zhang, Cheng Li, Austin Rovinski, Arjun Khurana, Ron Dreslinski, Trevor Mudge, Vinicius Petrucci, Lingjia Tang, and Jason Mars. Sirius: An Open End-to-End Voice and Vision Personal Assistant and Its Implications for Future Warehouse Scale Computers. In *Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, ASPLOS '15, New York, NY, USA, 2015. ACM. Acceptance Rate: 17%.