



Department for  
Science, Innovation  
& Technology

# Dull o reoleiddio deallusrwydd artiffisial sydd o blaid arloesi

Ymateb y Llywodraeth i'r ymgynghoriad

Chwefror 2024



# Dull o reoleiddio deallusrwydd artiffisial sydd o blaid arloesi: ymateb y Llywodraeth i'r ymgynghoriad

Cyflwynwyd i Senedd y DU gan yr Ysgrifennydd Gwladol dros Wyddoniaeth,  
Arloesedd a Thechnoleg drwy Orchymyn Ei Fawrhydi

Chwefror 2024



© Hawlfraint y Goron 2024

Trwyddedir y cyhoeddiad hwn o dan delerau Trwydded Llywodraeth Agored f3.0 oni nodir fel arall. Er mwyn gweld y drwydded hon, ewch i [nationalarchives.gov.uk/doc/open-government-licence/version/3](https://nationalarchives.gov.uk/doc/open-government-licence/version/3) neu ysgrifennwch at The Information Policy Team, The National Archives, Kew, London TW9 4DU, neu anfonwch neges e-bost i: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Lle rydym wedi nodi unrhyw wybodaeth am hawlfraint trydydd parti, bydd angen i chi gael caniatâd deiliaid yr hawlfraint dan sylw.

Mae'r cyhoeddiad hwn ar gael yn [www.gov.uk/dsit](http://www.gov.uk/dsit)

Dylid anfon unrhyw ymholiadau ynglŷn â'r cyhoeddiad hwn atom yn: [alt.formats@dsit.gov.uk](mailto:alt.formats@dsit.gov.uk)

SBN 978-1-5286-4565-2  
E03019481 02/24

Argraffwyd ar bapur sy'n cynnwys o leiaf 40% o ddeunydd ffeibr wedi'i ailgylchu

Argraffwyd yn y DU gan HH Associates Ltd. ar ran Rheolwr Llyfrfa Ei Fawrhydi

# Cynnwys

<b>1. Rhagair gan y Gweinidog</b>	<b>3</b>
<b>2. Crynodeb gweithredol</b>	<b>6</b>
<b>3. Rhestr termau</b>	<b>9</b>
<b>4. Cyflwyniad</b>	<b>12</b>
<b>5. Fframwaith rheoleiddio sy'n datblygu'n gyfochrog â thechnoleg sy'n datblygu'n gyflym</b>	<b>15</b>
5.1. Dull cymesur o reoleiddio'r defnydd o ddeallusrwydd artiffisial sy'n seiliedig ar gyd-destun	15
5.2. Ystyried yr achos dros gyfrifoldebau newydd i ddatblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog	29
5.3. Gweithio gyda phartneriaid rhyngwladol i hyrwyddo cydweithrediad effeithiol ar lywodraethu deallusrwydd artiffisial	40
5.4. Map ffordd o'n camau nesaf o ran rheoleiddio deallusrwydd artiffisial	45
<b>6. Crynodeb o dystiolaeth yr ymgynghoriad ac ymateb y llywodraeth</b>	<b>48</b>
6.1. Yr egwyddorion deallusrwydd artiffisial traws-sector diwygiedig	48
6.2. Dyletswydd statudol i roi sylw	51
6.3. Swyddogaethau canolog newydd i gefnogi'r fframwaith	53
6.4. Monitro a gwerthuso'r fframwaith	56
6.5. Galluoedd rheoleiddwyr	57
6.6. Adnoddau ar gyfer deallusrwydd artiffisial dibynadwy	59
6.7. Syniadau i gloi	62
6.8. Cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial	64
6.9. Modelau sylfaen a'r fframwaith rheoleiddio	67
6.10. Pyllau tywod a meiciau arbrofi deallusrwydd artiffisial	69
<b>Atodiad A: Dull ac ymgysylltu</b>	<b>72</b>
<b>Atodiad B: Rhestr o'r ymatebwyr i'r ymgynghoriad</b>	<b>80</b>
<b>Atodiad C: Crynodebau o gwestiynau unigol</b>	<b>84</b>
<b>Atodiad D: Crynodeb o dystiolaeth asesu effaith</b>	<b>102</b>

Dull o reoleiddio deallusrwydd artiffisial sydd o blaid arloesi: ymateb y Llywodraeth i'r ymgynghoriad

# 1. Rhagair gan y Gweinidog



Y Gwir Anrhydeddus Michelle Donelan AS, yr Ysgrifennydd Gwladol dros Wyddoniaeth, Arloesedd a Thechnoleg

Mae'r byd ar drothwy oes newydd hynod a lywir gan ddatblygiadau ym maes deallusrwydd artiffisial. Rwy'n credu bod y gwelliannau cyflym o ran galluedd deallusrwydd artiffisial yn gyfle sy'n codi unwaith mewn cenhedlaeth i bobl y DU chwyldroi ein gwasanaethau cyhoeddus er gwell a sicrhau canlyniadau hirdymor, diriaethol gwirioneddol i'n gwlad.

Rhagwelir y bydd marchnad deallusrwydd artiffisial y DU yn tyfu i fwy na \$1 triliwn (Doleri'r UD) erbyn 2035<sup>1</sup> – gan ddatgloi popeth o sgiliau a swyddi newydd i driniaethau sy'n achub bywydau ar gyfer clefydau creulon fel canser a dementia a fyddai wedi bod yn amhosibl eu dychmygu gynt. Fy uchelgais yw ein bod yn chwyldroi'r ffordd rydym yn darparu gwasanaethau cyhoeddus drwy ddod yn arweinydd byd-eang ar ddatblygu a chyflwyno deallusrwydd artiffisial yn ddiogel.

Rydym wedi gwneud mwy nag unrhyw lywodraeth erioed i wireddu hynny, ac mae ein cynllun yn gweithio. Y llynedd, gwnaethom gynnal yr Uwchgynhadledd gyntaf yn y byd ar Ddiogelwch Deallusrwydd Artiffisial, gan ddwyn ynghyd ddiwydiant, y byd academiaidd, a chymdeithas sifil ar y cyd â 28 o'r gwledydd blaenllaw ym maes deallusrwydd artiffisial a'r UE i gytuno ar Ddatganiad Bletchley – ymrwymiad pwysig i rannu cyfrifoldeb am liniaru risgiau deallusrwydd artiffisial ffin eithaf, cydweithio ar ddiogelwch ac ymchwil, a hyrwyddo ei botensial fel grym er da yn y byd sydd ohoni.

---

<sup>1</sup> [United Kingdom Artificial Intelligence Market](#), Gweinyddiaeth Masnach Ryngwladol yr UD, 2023.

Ni oedd y llywodraeth gyntaf yn y byd i gyhoeddi'n ffurfiol ein hasesiad o'r galluoedd a'r risgiau sy'n deillio o ddeallusrwydd artiffisial datblygedig. Gwnaeth adroddiadau a lywir gan ymchwil gan yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg a Swyddfa'r Llywodraeth dros Wyddoniaeth<sup>2</sup> osod y sylfeini dros gytundeb rhyngwladol ar werthuso'r sail wyddonol dros ddiogelwch deallusrwydd artiffisial.

Gwnaethom ddwyn ynghyd gonsortiw m pwerus o arbenigwyr yn ein Sefydliad Diogelwch Deallusrwydd Artiffisial. Dyma sefydliad a gefnogir gan y llywodraeth sy'n ymrwymedig i hyrwyddo diogelwch deallusrwydd artiffisial er budd y cyhoedd, sef y sefydliad cyntaf o'i fath yn y byd.

Yn sgil cyhoeddi ein papur gwyn ar reoleiddio deallusrwydd artiffisial ym mis Mawrth 2023, roeddwn am fabwysiadu dull gweithredu beiddgar ac ystyriol sy'n gadarn o blaid arloesi ac yn gadarn o blaid diogelwch. Roeddwn yn sylweddoli bod yn rhaid i'n dull gweithredu aros yn ddigon ystywyth i ymdrin â chyflymder digynsail datblygiadau, a hefyd yn ddigon cadarn ym mhob sector i fynd i'r afael â'r prif bryderon sy'n codi o ran niweidiau cymdeithasol posibl, risgiau camddefnyddio, a risgiau i ymreolaeth y mae ein hymarferion meddwl i arweinwyr wedi eu datgelu.

Mae'r dull gweithredu ystywyth hwn, sy'n seiliedig ar sectorau, wedi grymuso rheoleiddwyr i greu mesurau pwrpasol sydd wedi'u teilwra i'r gwahanol anghenion a risgiau a achosir gan wahanol rannau o'n heconomi. Gwnaeth y papur gwyn gynnig pum egwyddor glir i reoleiddwyr presennol y DU eu dilyn, gan nodi ein disgwyliadau o ran arloesi'n gyfrifol ym maes deallusrwydd artiffisial.

Mae'r dull ymarferol, synnwyr cyffredin hwn wedi cael ei groesawu a'i gymeradwyo gan y cwmnïau sydd ar y ffin eithaf o ran datblygu deallusrwydd artiffisial ac arbenigwyr diogelwch deallusrwydd artiffisial blaenllaw. Cafodd dull gweithredu'r DU ei gefnogi gan Google DeepMind, Microsoft, OpenAI ac Anthropic i gyd, a hefyd gan yr egin fusnesau deallusrwydd artiffisial ym Mhrydain, a llawer o leisiau blaenllaw yn y byd academaidd a chymdeithas sifil.

Wrth ystyried ein hymateb i'r ymgynghoriad, rwyf wedi ceisio manteisio ar y llwyddiant hwn a sbarduno ein cynlluniau i sicrhau mai Prydain yw'r lle mwyaf diogel ac arloesol i ddatblygu a chyflwyno deallusrwydd artiffisial yn y byd, gyda mwy na £100 miliwn i gefnogi arloesedd ym maes deallusrwydd artiffisial a rheoleiddio deallusrwydd artiffisial. Gan adeiladu ar adborth o'r ymgynghoriad, rydym wedi sefydlu swyddogaeth ganolog i sicrhau natur gydlynol ein dull rheoleiddio ym mhob rhan o'r llywodraeth, gan gynnwys drwy recriwtio tîm amlddisgyblaethol newydd i wneud gwaith asesu risg a monitro traws-sector rhag risgiau deallusrwydd artiffisial sy'n bodoli eisoes ac sy'n dod i'r amlwg.

Gyda chymorth y Fforwm Cydweithredu Rheoleiddio Digidol, rydym wedi lansio'r Hwb Deallusrwydd Artiffisial a Digidol, sef cynllun peilot ar gyfer gwasanaeth cynghori newydd sbon i gefnogi arloesedd sy'n cael ei redeg gan reoleiddwyr arbenigol, gan gynnwys y Swyddfa Gyfathrebiadau (Ofcom), yr Awdurdod Cystadleuaeth a Marchnadoedd, yr Awdurdod Ymddygiad Ariannol a Swyddfa'r Comisiynydd Gwybodaeth.<sup>3</sup> Rydym hefyd yn buddsoddi mewn cymorth newydd i reoleiddwyr feithrin eu harbenigedd technegol ymarferol ac yn cefnogi lansiad naw hwb ymchwil newydd ledled y DU i fanteisio ar rym deallusrwydd artiffisial ym mhopeth o fathemateg i ofal iechyd.

---

<sup>2</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>3</sup> [New advisory service to help businesses launch AI and digital innovations](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

Gan ddatblygu ein harweinwyr meddwl ar ddiogelwch, rydym hefyd yn amlinellu'r achos dros set o ofynion wedi'u targedu sy'n gyfrwymol ar ddatblygwyr modelau deallusrwydd artiffisial diben cyffredinol tra galluog yn y dyfodol er mwyn sicrhau bod deallusrwydd artiffisial soffistigedig a phwerus yn datblygu mewn ffordd sy'n ddiogel. A bydd ein hymgyngoriadau wedi'u targedu ar ein cofrestr risgiau deallusrwydd artiffisial draws-economi a'n fframwaith monitro a gwerthuso yn ymgysylltu â lleisiau blaenllaw ymhlith rheoleiddwyr, y byd academiaidd, cymdeithas sifil, a'r diwydiant.

Bydd gan arbenigwyr technegol y Sefydliad Diogelwch Deallusrwydd Artiffisial rôl hollbwysig i'w chwarae wrth i ni ddatblygu ein dull o reoleiddio systemau diben cyffredinol tra galluog. Byddwn yn gweithio'n agos gyda datblygwyr deallusrwydd artiffisial, academyddion ac aelodau o gymdeithas sifil sy'n gallu cynnig safbwyntiau arbenigol annibynnol, a hefyd gyda'n partneriaid rhyngwladol cyn yr Uwchgynadleddau nesaf ar Ddiogelwch Deallusrwydd Artiffisial yng Ngweriniaeth Korea a Ffrainc.

Yn olaf, mae fy syniadau ynglŷn â rôl y DU i arwain ym maes deallusrwydd artiffisial yn mynd ymhell y tu hwnt i'r gorwel o'n blaenau ar hyn o bryd. Bydd angen i ni arwain ym maes ymchwil a fydd yn ein helpu i greu cymdeithas fwy gwydn sy'n barod am fyd lle mae technoleg deallusrwydd artiffisial uwch a'r modd i'w datblygu yn hygyrch yn gyffredinol. Mae hynny'n golygu gwella ein gallu i amddiffyn ein hunain rhag y rhai â bwriadau drwg a fydd yn ceisio defnyddio deallusrwydd artiffisial i wneud niwed a dylunio seilwaith newydd i'r rhyngwrwd ar gyfer byd digidol llawn systemau deallusrwydd artiffisial â phŵer i weithredu, ac mae hefyd yn golygu manteisio ar ddeallusrwydd artiffisial i wella agweddau allweddol ar ein cymdeithas megis trafodaethau democrataidd a chonsensws. Gall deallusrwydd artiffisial aros yn rym er budd y cyhoedd a rhaid iddo wneud hynny, a byddwn yn sicrhau bod hyn yn wir wrth i ni ddatblygu ein polisi yn y maes hwn.

Mae'r papur ymateb hwn yn gam penderfynol a chdir arall ymlaen o ran uchelgais y DU i arwain ym maes deallusrwydd artiffisial diogel a bod yn Oruwchbŵer Gwyddoniaeth a Thechnoleg erbyn diwedd y degawd. P'un a ydych yn datblygu deallusrwydd artiffisial, ei ddefnyddio, yn ymchwilio i'w ddiogelwch neu'n cynrychioli cymdeithas sifil, mae gennym oll ddiddordeb mewn gwireddu'r cyfleoedd sy'n gysylltiedig â datblygu deallusrwydd artiffisial diogel. Rwyf innau wedi fy ysbrydoli gan genhadaeth i wella bywydau pobl y DU drwy dechnoleg ac arloesedd, ac mae ein papur ymateb yn nodi sut yn union y caiff y genhadaeth honno ei gwireddu.



## 2. Crynodeb gweithredol

- Bu cyflymder cynnydd ym maes deallusrwydd artiffisial yn wahanol i bob technoleg flaenorol, ac mae'r manteision eisoes yn cael eu gwireddu ledled y DU: mae deallusrwydd artiffisial yn helpu i wneud ein swyddi yn fwy diogel a boddhaus, gwarchod ein bywyd gwyllt a gwrthsefyll newid yn yr hinsawdd, a gwneud ein gwasanaethau cyhoeddus yn fwy effeithlon. Nid yn unig y mae angen i ni gynllunio ar gyfer galluoedd a defnyddiau'r systemau deallusrwydd artiffisial sydd gennym heddiw, ond mae'n rhaid i ni hefyd baratoi ar gyfer dyfodol agos lle mae'r systemau mwyaf pwerus yn hygyrch yn gyffredinol ac yn llawer mwy galluog.<sup>4</sup>
- Mae'r DU yn arwain y byd o ran sut i ymateb i'r her hon. Wrth baratoi ar gyfer y fath ddyfodol rydym yn gadarn o blaid arloesi. Er mwyn gwireddu manteision anferth y technolegau hyn, mae'n rhaid i ni sicrhau dibynadwyedd deallusrwydd artiffisial a bod y cyhoedd yn ei fabwysiadu drwy ddull gweithredu sy'n gadarn o blaid diogelwch. Fel y nododd Prif Weinidog y DU mewn araith bwysig ym mis Hydref 2023, "the future of AI is safe AI. And by making the UK a global leader in safe AI, we will attract even more of the new jobs and investment that will come from this new wave of technology".<sup>5</sup> Er mwyn cyflawni hyn, mae'r DU yn buddsoddi mwy mewn deallusrwydd artiffisial nag unrhyw wlad arall yn y byd. Heddiw, rydym yn cyhoeddi dros £100 miliwn i helpu i wireddu datblygiadau arloesol newydd ym maes deallusrwydd artiffisial a chefnogi galluoedd technegol rheoleiddwyr.
- Mae ein fframwaith rheoleiddio yn adeiladu ar gryfderau presennol ein diwydiant deallusrwydd artiffisial sy'n ffynnu a'n hecosystem reoleiddio arbenigol. Rydym yn canolbwyntio ar sicrhau bod rheoleiddwyr yn barod i wynebu'r heriau a'r cyfleoedd newydd a all ddod i'w rhan o ganlyniad i ddeallusrwydd artiffisial. Drwy weithio'n agos gyda rheoleiddwyr er mwyn sicrhau cydlyniant drwyddi draw, rydym yn sicrhau y gall arloeswyr gyflwyno cynhyrchion newydd ar y farchnad yn ddiogel ac yn gyflym. Heddiw, rydym yn cyhoeddi sawl menter newydd i wneud y DU yn lle hyd yn oed yn well i adeiladu a defnyddio deallusrwydd artiffisial, gan gynnwys £10 miliwn i sbarduno galluoedd deallusrwydd artiffisial rheoleiddwyr; ymrwymiad newydd gan Ymchwil ac Arloesi yn y DU y bydd yn denu buddsoddiadau mewn ymchwil i ddeallusrwydd artiffisial yn y dyfodol er mwyn cefnogi sgiliau ac arbenigedd rheoleiddwyr; a phartneriaeth gwerth £9 miliwn â'r UD ynglŷn â deallusrwydd artiffisial cyfrifol fel rhan o'n Cronfa Partneriaethau Gwyddoniaeth Rhyngwladol.<sup>6</sup> Drwy hyn a gwaith arall ym mhob rhan o'r llywodraeth, bydd y DU yn parhau i ymateb i risgiau mewn ffordd gymesur ac effeithiol, sy'n ymdrechu i arwain syniadaeth deallusrwydd artiffisial yn y blynyddoedd sydd i ddod.
- Ym mis Mawrth 2023, gwnaethom gyhoeddi ein papur gwyn ar reoleiddio deallusrwydd artiffisial, gan nodi ein cynigion cychwynnol i ddatblygu fframwaith rheoleiddio ar gyfer deallusrwydd artiffisial sydd o blaid arloesi. Amlinellodd y fframwaith arfaethedig bum egwyddor traws-sector ar gyfer rheoleiddwyr presennol y DU i'w dehongli a'u cymhwyso o fewn eu cylchoedd gwaith. Gwnaethom hefyd

---

<sup>4</sup> Er mwyn cefnogi gwaith cynllunio a datblygu polisiau'r llywodraeth, ac yn wyneb cryn ansicrwydd, mae Swyddfa'r Llywodraeth dros Wyddoniaeth wedi paratoi adroddiad rhagargoeli sy'n amlinellu senarios posibl a all godi yng nghyd-destun datblygu a lledaenu deallusrwydd artiffisial a'i effaith yn 2030. Gweler: [Future risks of frontier AI \(Annex A\)](#), Swyddfa'r Llywodraeth dros Wyddoniaeth, 2023. Caiff adroddiad llawn ar y senarios ei gyhoeddi cyn hir (ni fydd yr adroddiad hwn yn datgan polisi'r llywodraeth).

<sup>5</sup> [Prime Minister's speech on AI: 26 October 2023](#), Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.

<sup>6</sup> [International Science Partnerships Fund \(ISPF\)](#), Ymchwil ac Arloesi yn y DU, 2023.

gynnig swyddogaeth ganolog newydd i sicrhau bod y gyfundrefn yn gydlynol a mynd i'r afael â bylchau ym maes rheoleiddio. Drwy'r dull rheoleiddio hyblyg ac addasol hwn rydym wedi llwyddo i weithredu'n benderfynol ac ymateb i gynnydd technolegol.

- Cafodd ein fframwaith seiliedig ar gyd-destun gefnogaeth fawr gan randdeiliaid ym mhob rhan o gymdeithas ac rydym wedi gweithredu'n ddi-oed i'w roi ar waith. Mae'n dda gennym nodi bod nifer o reoleiddwyr eisoes yn cymryd camau yn unol â'n dull gweithredu arfaethedig, o adolygiad o fodolau sylfaen yr Awdurdod Ymddygiad Ariannol i'r canllawiau wedi'u diweddarau ar ddiogelu data a deallusrwydd artiffisial gan Swyddfa'r Comisiynydd Gwybodaeth. Rydym yn gofyn i nifer o reoleiddwyr gyhoeddi diweddiariad yn amlinellu eu dull gweithredu strategol o ymdrin â deallusrwydd artiffisial erbyn 30 Ebrill 2024.
- Rydym eisoes wedi dechrau datblygu'r swyddogaeth ganolog i gefnogi trefniadau effeithiol ar gyfer monitro risg, cydgysylltu rheolaidd, a chyfnewid gwybodaeth. Bydd ein pecyn newydd gwerth £10 miliwn i roi hwb i alluoedd deallusrwydd artiffisial rheoleiddwyr, a grybwyllir uchod, yn helpu ein rheoleiddwyr i ddatblygu adnoddau ymchwil ac ymarferol arloesol er mwyn adeiladu sylfeini eu harbenigrwydd mewn deallusrwydd artiffisial a'u gallu i fynd i'r afael â'i risgiau bob dydd yn eu priod feysydd. Heddiw, rydym hefyd yn cyhoeddi canllawiau newydd i helpu rheoleiddwyr i weithredu'r egwyddorion yn effeithiol ac mae'r Fforwm Cydweithredu Rheoleiddio Digidol yn rhannu manylion am y meini prawf cymhwysra i gael y cymorth a gynigir o dan gynllun peilot deallusrwydd artiffisial a'r Hwb Digidol.
- Rydym yn cefnogi'r dull gweithredu hwn drwy roi cymorth ehangach ar gyfer yr ecosystem deallusrwydd artiffisial, gan gynnwys ymrwymo i ddarparu dros £1.5 biliwn yn 2023 i adeiladu'r genhedlaeth nesaf o uwchgyfrifiaduron yn y sector cyhoeddus a heddiw rydym yn cyhoeddi swm o £80 miliwn i hybu ymchwil ym maes deallusrwydd artiffisial drwy lansio naw hwb ymchwil newydd ledled y DU i sbarduno datblygiadau arloesol trawsnewidiol. Ym mis Tachwedd 2023, gwnaeth y Prif Weinidog ddwyn ynghyd y rhai blaenllaw ym maes deallusrwydd artiffisial o bedwar ban byd ar gyfer yr Uwchgynhadledd gyntaf ar Ddiogelwch Deallusrwydd Artiffisial lle y trafodwyd ac y cytunwyd ar gamau gweithredu i ymdrin â'r risgiau sy'n dod i'r amlwg o ganlyniad i ddatblygu a chyflwyno'r systemau deallusrwydd artiffisial mwyaf pwerus. Nododd datblygwyr blaenllaw ym maes deallusrwydd artiffisial y camau y maent eisoes yn eu cymryd i wneud modelau yn ddiogel a gwnaethant ymrwymo i rannu'r modelau deallusrwydd artiffisial mwyaf pwerus â llywodraethau i'w profi fel y gallwn sicrhau diogelwch heddiw a pharatoi ar gyfer risgiau yfory.
- Ein cyfraniad technegol cychwynnol at yr ymdrech ryngwladol hon yw drwy greu Sefydliad Diogelwch Deallusrwydd Artiffisial i arwain gwerthusiadau ac ymchwil i ddiogelwch yn Llywodraeth y DU, a chydweithio â phartneriaid ledled y byd, gan gynnwys yn yr UD. Tanlinellodd yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial natur fyd-eang datblygu a chyflwyno deallusrwydd artiffisial, gan ddangos bod angen gwneud rhagor o waith i greu dull cydlynol a chydweithredol o lywodraethu rhyngwladol.
- Mae ein dull gweithredu cyffredinol – sy'n cyfuno egwyddorion traws-sector a fframwaith cyd-destun penodol, arweinyddiaeth a chydweithio rhyngwladol, a mesurau gwirfoddol ar ddatblygwyr – yn briodol heddiw am ei fod yn ei gwneud yn bosibl i ni ymateb i ddatblygiadau cyflym ac ansicr ym maes deallusrwydd artiffisial. Fodd bynnag, bydd yr heriau a ddaw yn sgil technolegau deallusrwydd artiffisial yn gofyn, yn y pen draw, am weithredu deddfwriaethol ym mhob gwlad unwaith y bydd ddealltwriaeth o risg yn fwy aeddfed. Yn y ddogfen hon, rydym yn adeiladu ar ein fframwaith o blaid arloesi a'n camau gweithredu o blaid diogelwch drwy nodi

ein syniadaeth gynnar a'r cwestiynau y bydd angen i ni eu hystyried yn ystod y cam nesaf o'n dull rheoleiddio. Gan gydnabod nad oes unrhyw atebion hawdd, byddwn yn gweithio'n agos gyda chymdeithas sifil, diwydiant, a phartneriaid rhyngwladol i archwilio'r materion hyn, a byddwn yn dryloyw o ran rhannu safbwyntiau arbenigol cynnar arnynt.

- Wrth i systemau deallusrwydd artiffisial ddatblygu o ran gallu a'u heffaith ar gymdeithas, mae'n amlwg y bydd angen rhai mesurau gorfodol yn y pen draw rhwng pob awdurdodaeth er mwyn mynd i'r afael â niweidiau posibl deallusrwydd artiffisial, sicrhau diogelwch y cyhoedd, a gadael i ni wireddu'r cyfleoedd trawsnewidiol y mae'r dechnoleg yn eu cynnig. Fodd bynnag, pe baem yn gweithredu cyn ein bod yn deall y risgiau a'r mesurau lliniaru priodol yn iawn byddai hynny'n amharu ar ein gallu i fanteisio ar gynnydd technolegol a hefyd yn golygu na allwn ymaddasu'n gyflym i'r risgiau sy'n dod i'r amlwg. Rydym yn mynd i gymryd ein hamser i gael y pethau hyn yn iawn – byddwn yn deddfu pan fyddwn yn hyderus mai dyna'r peth cywir i'w wneud.
- Rydym wedi rhoi pwyslais penodol ar yr heriau y mae systemau deallusrwydd artiffisial diben cyffredinol tra galluog yn eu hachosi i fframwaith seiliedig ar gyd-destun. Yma, rydym yn amlinellu achos o blaid arloesi ar gyfer gofynion cyfrwymol, wedi'u targedu pellach ar nifer bach o sefydliadau sy'n datblygu systemau deallusrwydd artiffisial diben cyffredinol tra galluog er mwyn sicrhau eu bod yn atebol am wneud y technolegau hyn yn ddigon diogel. Gellir gwneud hyn a hefyd ganiatáu i'n rheoleiddwyr arbenigol lunio rheolau effeithiol ar gyfer defnyddio deallusrwydd artiffisial yn eu cylchoedd gwaith.
- Yn y misoedd nesaf, byddwn yn pennu'n ffurfiol ein gweithgareddau i gefnogi galluoedd a threfniadau cydgysylltu rheoleiddwyr, gan gynnwys pwyllgor llywio newydd a fydd yn cynnwys cynrychiolwyr y llywodraeth a rheoleiddwyr er mwyn helpu i gydlynu tirwedd rheoleiddio deallusrwydd artiffisial. Byddwn yn cynnal ymgynghoriadau wedi'u targedu ar ein cofrestr risgiau deallusrwydd artiffisial draws-economi ac yn cynllunio i asesu'r fframwaith rheoleiddio. Byddwn yn parhau â'n gwaith i fynd i'r afael â'r materion allweddol a wynebwn heddiw, o ymyrryd ag etholiadau i wahaniaethu i gyfraith eiddo deallus, a risgiau mwyaf dybryd yfory, megis bioddiogelwch a chysoni deallusrwydd artiffisial. Byddwn hefyd yn parhau i arwain sgysiau rhyngwladol ar lywodraethu deallusrwydd artiffisial mewn nifer o fforymau a mentrau cyn yr Uwchgynadleddau nesaf ar Ddiogelwch Deallusrwydd Artiffisial yng Ngweriniaeth Korea a Ffrainc.

## 3. Rhestr termau

**Awtonomaidd:** Yn gallu gweithredu, cymryd camau, neu wneud penderfyniadau heb fwriad penodol na goruchwyliaeth ddynol.

**Cyflwynwyr deallusrwydd artiffisial:** Unrhyw unigolyn neu sefydliad sy'n cyflenwi neu sy'n defnyddio cymhwysiad deallusrwydd artiffisial i ddarparu cynnyrch neu wasanaeth i ddefnyddiwr terfynol.

**Cyfrifiannu:** Pŵer prosesu cyfrifiannol, gan gynnwys Unedau Prosesu Canolog, Unedau Prosesu Graffeg, a mathau eraill o galedwedd, a ddefnyddir i redeg modelau ac algorithmau deallusrwydd artiffisial.

**Cylch oes deallusrwydd artiffisial:** Pob digwyddiad a phroses sy'n ymwneud ag oes system deallusrwydd artiffisial, o'i dechrau hyd at ei datgomiynu, gan gynnwys dylunio, ymchwil, hyfforddi, datblygu, cyflwyno, integreiddio, gweithredu, cynnal a chadw, gwerthu, defnyddio, a llywodraethu.

**Cysoni:** Y broses o sicrhau bod nodau ac ymddygiadau system deallusrwydd artiffisial yn unol â gwerthoedd a bwriadau dynol.

**Datblygu systemau diben cyffredinol tra galluog:** Is-adran o ddatblygwyr deallusrwydd artiffisial. Mae'r sefydliadau hyn yn buddsoddi llawer o adnoddau i ddylunio, adeiladu, a rhaghyfforddi'r modelau sylfaen deallusrwydd artiffisial mwyaf galluog. Gall y modelau hyn fod yn sail i amrywiaeth eang o gymwysiadau deallusrwydd artiffisial a gallant gael eu cyflwyno'n uniongyrchol neu eu haddasu gan ddatblygwyr deallusrwydd artiffisial i lawr y gadwyn.

**Datblygwyr deallusrwydd artiffisial:** Sefydliadau neu unigolion sy'n dylunio, yn adeiladu, yn hyfforddi, yn addasu, neu'n cyfuno modelau a chymwysiadau deallusrwydd artiffisial.

**Deallusrwydd Cyffredinol Artiffisial (AGI):** Ffurf ddamcaniaethol ar ddeallusrwydd artiffisial uwch a fyddai'n meddu ar alluoedd sy'n debyg i alluoedd bodau dynol neu sy'n rhagori arnynt yn y gwaith mwyaf gwerthfawr yn economaidd.<sup>7</sup> mae nifer o gwmnïau deallusrwydd artiffisial wedi nodi'n gyhoeddus mai eu nod yw adeiladu deallusrwydd cyffredinol artiffisial a'u bod yn credu y gellir cyflawni hynny o fewn yr ugain mlynedd nesaf. Cred arbenigwyr eraill na fydd modd i ni adeiladu deallusrwydd cyffredinol artiffisial am ddegawdau lawer, os byth.

**Defnyddwyr terfynol deallusrwydd artiffisial:** Unrhyw unigolyn neu sefydliad arfaethedig neu wirioneddol sy'n defnyddio cynnyrch neu wasanaeth sy'n seiliedig ar ddeallusrwydd artiffisial wrth iddo gael ei gyflwyno.

**Galluoedd:** Yr ystod o dasgau neu swyddogaethau y gall system deallusrwydd artiffisial eu cyflawni a'u medrusrwydd i'w cyflawni.

**Gweithredyddion deallusrwydd artiffisial:** systemau deallusrwydd artiffisial awtonomaidd sy'n cyflawni camau dilyniannol lluosog – weithiau gan gynnwys camau fel pori'r rhyngwrwd, anfon negeseuon e-bost, neu anfon cyfarwyddiadau at gyfarpar ffisegol – er mwyn ceisio cwblhau tasg neu nod lefel uchel.

**Risgiau deallusrwydd artiffisial:** Y canlyniadau negyddol neu niweidiol posibl sy'n deillio o ddatblygu neu gyflwyno systemau deallusrwydd artiffisial.

---

<sup>7</sup> [How should AI systems behave, and who should decide?](#), OpenAI, 2023.

**Rhyngwyneb Rhaglennu Cymwysiadau** : Set o reolau a phrotocolau sy'n ei gwneud yn bosibl i systemau deallusrwydd artiffisial a chymwysiadau meddalwedd eraill integreiddio a chyfathrebu â'i gilydd.

**Twyllwbyodaeth**: Gwybodaeth sy'n fwriadol anwir sy'n cael ei lledaenu gyda'r bwriad o dwyllo neu gamarwain.

**Ymaddasolrwydd**: Y gallu i weld patrymau a gwneud penderfyniadau mewn ffyrdd nas rhagwelwyd yn uniongyrchol gan raglenwyr dynol.

**Modelau sylfaen**: Modelau dysgu peirianyddol sy'n cael eu hyfforddi ar lawer iawn o ddata y gellir eu haddasu at amrywiaeth eang o dasgau.

**Deallusrwydd artiffisial ffin eithaf**: Ar gyfer yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, gwnaethom ddiffinio deallusrwydd artiffisial ffin eithaf fel modelau sy'n gallu cyflawni amrywiaeth eang o dasgau sydd â galluoedd cystal â'r modelau mwyaf datblygedig a welir heddiw neu sydd â galluoedd sy'n rhagori arnynt. Yn y papur hwn, rydym yn canolbwyntio ar ddatblygwyr modelau deallusrwydd artiffisial diben cyffredinol tra galluog fel targed ein cynigion ar gyfer cyfrifoldebau newydd.

**Camwbyodaeth**: Gwybodaeth anghywir neu gamarweiniol sy'n cael ei lledaenu heb fwriad niweidiol.

**Diogelwch**: Diogelwch, llesiant ac awtonomi cymdeithas sifil a'r boblogaeth.<sup>8</sup> Yn y cyhoeddiad hwn, defnyddir diogelwch yn aml i ddisgrifio atal niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu ddiogelu rhagddynt. Mae diogelwch deallusrwydd artiffisial yn cyfeirio at ddiogelu systemau deallusrwydd artiffisial rhag ymyrraeth dechnegol megis seiberymosodiadau.<sup>9</sup>

**Perfformiad goruwchddynol**: Pan fydd model deallusrwydd artiffisial yn dangos galluoedd sy'n rhagori ar feincnodau gallu dynol ar gyfer tasg neu weithgarwch penodol.

## Blwch 1: Gwahanol fathau o systemau deallusrwydd artiffisial

Yn ein papur trafod ar alluoedd a risgiau deallusrwydd artiffisial ffin eithaf,<sup>10</sup> nodwyd gennym fod diffiniadau o ddeallusrwydd artiffisial yn aml yn heriol oherwydd datblygiadau cyflym yn y dechnoleg.

At ddibenion datblygu dull rheoleiddio cymesur sy'n mynd i'r afael mewn ffordd effeithiol â'r risgiau sy'n deillio o'r systemau deallusrwydd artiffisial mwyaf pwerus, ar hyn o bryd rydym yn gwahaniaethu rhwng:

1. **deallusrwydd artiffisial diben cyffredinol tra galluog**: Modelau sylfaen sy'n gallu cyflawni amrywiaeth eang o dasgau sydd â galluoedd cystal â'r modelau mwyaf datblygedig a welir heddiw neu sydd â galluoedd sy'n rhagori arnynt. Yn gyffredinol, bydd modelau o'r fath yn rhychwantu galluoedd newydd hyd at alluoedd arbenigol gyda rhai hyd yn oed yn dangos perfformiad goruwchddynol mewn nifer o dasgau.

<sup>8</sup> [Safety and Security Risks of Generative Artificial Intelligence to 2025](#), Swyddfa'r Llywodraeth dros Wyddoniaeth, 2023.

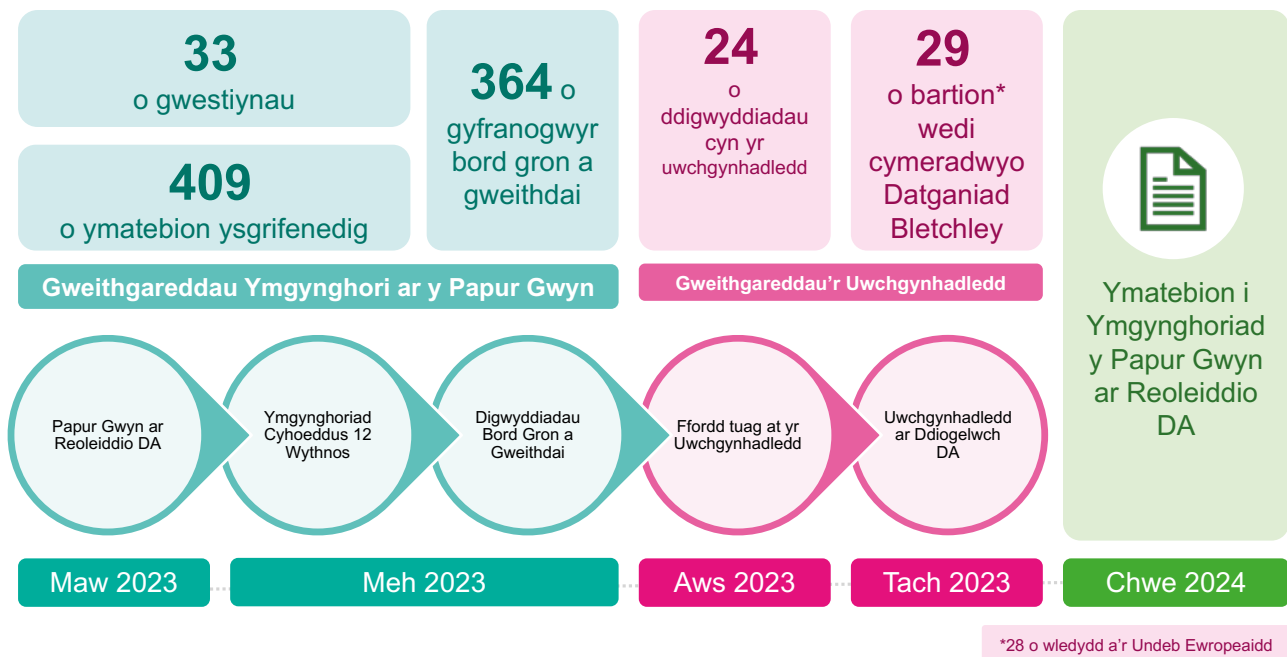
<sup>9</sup> Rhown ragor o fanylion ar y maes hwn fel rhan o'n disgrifiad o'r egwyddor diogelwch a chadernid traws-sector yn y papur gwyn ar reoleiddio deallusrwydd artiffisial. Gweler: [AI regulation: a pro-innovation approach](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>10</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

2. **Deallusrwydd artiffisial tra galluog:** Modelau sylfaen sy'n gallu cyflawni set gul o dasgau, fel arfer mewn maes penodol megis bioleg, sydd â galluoedd cystal â'r modelau mwyaf datblygedig a welir heddiw neu sydd â galluoedd sy'n rhagori arnynt. Yn gyffredinol, bydd modelau o'r fath yn dangos galluoedd goruwchddynol ar y tasgau neu'r parthau cul hyn.
3. **Deallusrwydd artiffisial â phŵer i weithredu neu weithredyddion deallusrwydd artiffisial:** Is-set ddatblygol o dechnolegau deallusrwydd artiffisial sy'n gallu cwblhau tasgau'n gymwys dros derfynau amser hir lle mae camau lluosog. Mae'r systemau hyn yn gallu defnyddio adnoddau megis amgylcheddau codio, y rhyngwyd, a modelau deallusrwydd artiffisial cul i gwblhau tasgau.

## 4. Cyflwyniad

1. Mae sector deallusrwydd artiffisial y DU yn ffynnu. Mae'r diwydiant deallusrwydd artiffisial yn y DU yn cyflogi dros 50,000 o bobl ac yn cyfrannu £3.7 biliwn i'n heconomi.<sup>11</sup> Mae ein prifysgolion yn datblygu rhai o'r ymchwilwyr gorau a'r dalent orau ym maes deallusrwydd artiffisial ac mae'r DU yn y trydydd safle ledled y byd o ran nifer y busnesau ungorn ac egin fusnesau sydd yma.<sup>12</sup>
2. Ein nod yw gwneud y DU yn lle gwych i adeiladu a datblygu deallusrwydd artiffisial sy'n newid ein bywydau er gwell. Deallusrwydd artiffisial yw'r dechnoleg sy'n diffinio ein hoes ac mae'r DU yn arwain y byd drwy ein hymateb.
3. Ym mis Mawrth 2023, gwnaethom gyhoeddi papur gwyn yn nodi ein cynigion i sefydlu fframwaith rheoleiddio ar gyfer deallusrwydd artiffisial er mwyn sbarduno arloesi mewn ffordd ddiogel a chyfrifol.<sup>13</sup> Pennwyd pum egwyddor i reoleiddwyr eu dehongli a'u cymhwyso yn eu priod feysydd. Gwnaethom hefyd gynnwys cynigion ar gyfer swyddogaeth ganolog yn y llywodraeth i gynnal nifer o weithgareddau megis asesu risg a chydgylltu trefniadau rheoleiddio i'n helpu i allu addasu a bod yn gydlynol.
4. Cynhaliwyd ymgynghoriad cyhoeddus dro 12 wythnos ar ein cynigion.<sup>14</sup> Rydym bellach wedi dadansoddi'r dystiolaeth (gweler Atodiad A am fanylion) sydd wedi llywio'r ffordd rydym yn gweithredu. Hoffem ddiolch i bawb am y sylwadau a gyflwynwyd. Rydym hefyd wedi ymgorffori yn ein hymateb y cyflawniadau allweddol o'r Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial ym mis Tachwedd 2023, yn ogystal â themâu o'n gweithgarwch ymgysylltu cyn yr Uwchgynhadledd.



<sup>11</sup> Mae cwmnïau deallusrwydd artiffisial penodol mawr yn gwneud cyfraniadd mawr i economi'r DU. Amcangyfrifir gwerth ychwanegol gros o £400k fesul cyflogai, sef mwy na dwbl yr amcangyfrifon cyffelyb ar gyfer cwmnïau penodol mawr mewn sectorau eraill. Gweler: [AI Sector Study 2022](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>12</sup> [The Global AI Index](#), Tortoise Media, 2023.

<sup>13</sup> [AI regulation: a pro-innovation approach](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>14</sup> [AI regulation: a pro-innovation approach – policy proposals](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

5. Mae datblygiadau ym maes deallusrwydd artiffisial yn parhau i gyflymu. Yn y cyfnod cyn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, gwnaethom gyhoeddi papur trafodaeth ar risgiau a galluoedd deallusrwydd artiffisial a oedd yn dangos bod y tueddiadau hyn yn debygol o barhau yn unol â'r cwmnïau sy'n creu'r technolegau hyn, drwy ddefnyddio mwy o gyfrifiannu, mwy o ddata ac algorithmau cynyddol effeithlon.<sup>15</sup> Mae rhai labordai deallusrwydd artiffisial ffin eithaf wedi dweud mai eu nod yw adeiladau systemau deallusrwydd artiffisial sy'n fwy galluog na bodau dynol wrth wneud nifer o dasgau.<sup>16</sup>

6. Mae galluoedd uwch yn creu cyfleoedd newydd. Mae deallusrwydd artiffisial eisoes yn newid y ffordd rydym yn byw ac yn gweithio. Mae gweithwyr sy'n defnyddio deallusrwydd artiffisial mewn sectorau sy'n amrywio o weithgynhyrchu i gyllid wedi nodi gwelliannau o ran mwynhad swydd, perfformiad ac iechyd.<sup>17</sup> Bydd deallusrwydd artiffisial yn newid y tasgau rydym yn eu gwneud a'r sgiliau sydd eu hangen arnom i'w cyflawni'n dda.<sup>18</sup> Mae datblygiadau deallusrwydd artiffisial diweddar hefyd yn newid y ffordd rydym yn treulio ein hamser hamdden, gyda systemau deallusrwydd artiffisial pwerus yn sail i'r sgwrsfotiau a'r cynhyrchwyr delweddau sydd wedi dod yn rhai o'r cymwysiadau sydd wedi gweld y twf mwyaf erioed ymhlith defnyddwyr.<sup>19</sup> Mae deallusrwydd artiffisial tra galluog eisoes yn trawsnewid sectorau, o'n helpu i warchod ein bywyd gwyllt<sup>20</sup> i newid y ffordd rydym yn nodi ac yn trin clefydau.<sup>21</sup>

7. Fodd bynnag, mae deallusrwydd artiffisial mwy pwerus hefyd yn codi risgiau newydd, mwy o faint. Er enghraifft, gall sgwrsfotiau deallusrwydd artiffisial wneud gwybodaeth anwir yn fwy amlwg<sup>22</sup> gall system deallusrwydd artiffisial dra galluog gael ei gamddefnyddio i gyflawni troseddau. Er enghraifft, gallai model sydd wedi cael ei ddylunio i ddarganfod cyffuriau gael ei ddefnyddio'n faleisus o bosibl i greu cyfansoddion niweidiol.<sup>23</sup>

8. Gall deallusrwydd artiffisial hefyd drawsnewid bywyd mewn ffyrdd sylfaenol sy'n anodd eu rhagweld. Er enghraifft, efallai y gall systemau deallusrwydd artiffisial â phŵer i weithredu yn y dyfodol anelu at nodau cymhleth heb fawr ddim goruchwyliaeth ddynol, sy'n codi cwestiynau ynglŷn â sut y gall gweithredwyr deallusrwydd artiffisial barhau i fod yn briodoladwy, sut maen nhw'n gofyn am gymeradwyaeth cyn gweithredu, a sut y gellir ymyrryd â nhw.

9. Mae technolegau deallusrwydd artiffisial yn achosi cryn ansicrwydd sy'n gofyn am ddull rheoleiddio ystywyth sy'n cefnogi arloesedd ond sydd hefyd yn addasu i ymdrin â risgiau newydd. Yn yr ymateb hwn i'r ymgynghoriad, rydym yn dangos sut mae ein dull gweithredu hyblyg eisoes yn ymdrin â risgiau allweddol sy'n gysylltiedig â deallusrwydd artiffisial a sut rydym yn cryfhau'r fframwaith hwn ymhellach (adran 5.1). Rydym hefyd yn nodi ein syniadau cychwynnol ar gyfrifoldeb newydd posibl ar ddatblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog ochr yn ochr â'r ymrwymadau gwirfoddol a sicrhawyd

<sup>15</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>16</sup> [Race towards 'autonomous' AI agents grips Silicon Valley](#), Anna Tong a Jeffrey Dastin, 2023; [Introducing superalignment](#), Jan Leike a Ilya Sutskever (OpenAI), 2023; [AI could be one of humanity's most useful inventions](#), Google Deepmind, dim dyddiad.

<sup>17</sup> [Employment Outlook 2023: artificial intelligence and jobs](#), y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, 2023.

<sup>18</sup> [Generative AI and the UK labour market](#), KPMG, 2023; [The economic potential of generative AI: the next productivity frontier](#), McKinsey, 2023; [What drives UK firms to adopt AI and robotics, and what are the consequences for jobs?](#), Sefydliad Dyfodol Gwaith, 2023.

<sup>19</sup> [ChatGPT is the fastest growing app in the history of web applications](#), Cindy Gordon, 2023.

<sup>20</sup> [Using AI to monitor trackside Britain's wildlife](#), y Gymdeithas Sŵoleg, Llundain, 2023.

<sup>21</sup> [A foundation model for generalizable disease detection from retinal images](#), Esmá Aimeur et al., 2023.

<sup>22</sup> [Synthetic lies: understanding AI-generated misinformation and evaluating algorithmic and human solutions](#), Jiawei Zhou et al., 2023; [Fake news, disinformation and misinformation in social media: a review](#), Yukun Zhou et al., 2023; [AI could create a perfect storm of climate misinformation](#), Victor Galaz et al., 2023.

<sup>23</sup> [Dual use of artificial-intelligence-powered drug discovery](#), Fabio Urbina et al., 2022.



Dull o reoleiddio deallusrwydd artiffisial sydd o blaid arloesi: ymateb y Llywodraeth i'r ymgynghoriad

yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial (adran 5.2). Yn adran 6, rydym yn rhoi crynodeb o'r dystiolaeth a gyflwynwyd i'n hymgyngoriad ochr yn ochr â'n ymateb ffurfiol.

## 5. Fframwaith rheoleiddio sy'n datblygu'n gyfochrog â thechnoleg sy'n datblygu'n gyflym

10. Yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, gwnaethom gynnig pum egwyddor draws-sector i reoleiddwyr presennol eu dehongli a'u cymhwysu yn eu priod gylchoedd gwaith er mwyn sbarduno arloesi'n ddiogel ac yn gyfrifol ym maes deallusrwydd artiffisial<sup>24</sup>, sef:

- Diogelwch a chadernid.
- Tryloywder ac esboniadwyedd priodol.
- Tegwch
- Atebolrwydd a llywodraethu.
- Y gallu i herio a cheisio iawn.

11. Rydym yn croesawu'r gefnogaeth gref i'r egwyddorion hyn drwy'r ymgynghoriad. Maent yn sylfaenol i'n dull gweithredu. Rydym yn dal yn ymrwymedig i weithredu'n seiliedig ar gyd-destun sy'n osgoi rheolau diwahân, diangen sy'n gymwys i bob technegol deallusrwydd artiffisial, ni waeth sut y caiff ei defnyddio. Dyma'r ffordd orau o sicrhau dull gweithredu ystywyth sy'n sefyll prawf amser.

12. Mae'n bleser gennym weld sut mae rheoleiddwyr eisoes yn gweithredu ein hegwyddorion yn annibynnol. Yn y papur gwyn, gwnaethom dynnu sylw at bwysigrwydd swyddogaeth ganolog i gefnogi galluedd a threfniadau cydgysylltu rheoleiddwyr. Rydym wedi gwneud cynnydd da o ran sefydlu'r swyddogaeth hon yn y llywodraeth. Nodwn isod sut rydym yn ei chryfhau ymhellach, gan gynnwys cyllid newydd, yn adran 5.1. Rydym hefyd yn dangos sut mae rheoleiddwyr a'r llywodraeth yn ymdrin â rhai o'r materion pwysicaf a wynebwn heddiw.

13. Yn adran 5.2, rydym yn nodi rhai o'r heriau rheoleiddio sy'n codi o ganlyniad i ddatblygiad cyflym systemau diben cyffredinol tra galluog; sut rydym yn mynd i'r afael â'r rhain drwy fesurau gwirfoddol, gan gynnwys y rhai y cytunwyd arnynt yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial; a pha gyfrifoldebau ychwanegol y bydd eu hangen yn y dyfodol i ymdrin â risgiau'n effeithiol.

### 5.1. Dull cymesur o reoleiddio'r defnydd o ddeallusrwydd artiffisial sy'n seiliedig ar gyd-destun

5.1.1. Mae rheoleiddwyr yn cymryd camau gweithredol yn unol â'r fframwaith.

14. Ers cyhoeddi'r papur gwyn ar reoleiddio deallusrwydd artiffisial, mae nifer o reoleiddwyr wedi amlinellu gwaith yn unol â'n dull gweithredu yn seiliedig ar egwyddorion. Er enghraifft, cyhoeddodd yr Awdurdod Cystadleuaeth a Marchnadoedd adolygiad o fodolau sylfaen er mwyn deall y cyfleoedd a'r risgiau o ran cystadleuaeth a diogelwch

---

<sup>24</sup> [AI regulation: a pro-innovation approach](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

defnyddwyr.<sup>25</sup> Diweddarodd Swyddfa'r Comisiynydd Gwybodaeth ganllawiau ynglŷn â sut mae cyfreithiau diogelu data yn gymwys i systemau deallusrwydd artiffisial er mwyn cynnwys tegwch.<sup>26</sup> Er mwyn sicrhau diogelwch deallusrwydd artiffisial, mae rheoleiddwyr megis y Swyddfa Marchnadoedd Nwy a Thrydan (Ofgem) a'r Awdurdod Hedfan Sifil yn gweithio ar strategaethau deallusrwydd artiffisial i'w cyhoeddi yn ddiweddarach eleni. Mae hyn yn adeiladu ar waith rheoleiddwyr a arweiniodd y ffordd o ran egluro sut mae fframweithiau presennol yn gymwys i risgiau deallusrwydd artiffisial yn eu priod feysydd, megis Rhaglen Newid Meddalwedd a Deallusrwydd Artiffisial fel Dyfais Feddygol 2021 yr Asiantaeth Rheoleiddio Meddyginiaethau a Chynhyrchion Gofal Iechyd ynglŷn â gofynion ar gyfer meddalwedd a deallusrwydd artiffisial a ddefnyddir mewn dyfeisiau meddygol.<sup>27</sup>

15. Mae'n bwysig bod y cyhoedd yn gallu gweld yn union sut mae rheoleiddwyr yn ymgorffori'r egwyddorion yn eu gwaith. Mae'r llywodraeth wedi ysgrifennu at nifer o reoleiddwyr y mae deallusrwydd artiffisial yn effeithio arnynt yn gofyn iddynt gyhoeddi diweddariad yn amlinellu eu dull gweithredu strategol o ymdrin â deallusrwydd artiffisial erbyn 30 Ebrill 2024.<sup>28</sup> Anogwn y rheoleiddwyr i gynnwys:

- Amlinelliad o'r camau y maent yn eu cymryd yn unol â'r disgwyliadau a nodwyd yn y papur gwyn.
- Dadansoddiad o'r risgiau sy'n gysylltiedig â deallusrwydd artiffisial yn y sectorau a'r gweithgareddau y maent yn eu rheoleiddio a'r camau y maent yn eu cymryd i ymdrin â'r rhain.
- Esboniad o'u gallu ar hyn o bryd i ymdrin â deallusrwydd artiffisial o gymharu â'u hasesiad o ofynion, a'r camau gweithredu y maent yn eu cymryd i sicrhau bod ganddynt y strwythurau a'r sgiliau cywir yn eu lle.
- Rhagolwg ar gynlluniau a gweithgareddau dros y 12 mis nesaf.

16. Pan gyhoeddwyd y papur gwyn ar reoleiddio deallusrwydd artiffisial gennym, cynigiwyd y byddai'r egwyddorion yn cael eu sefydlu ar sail anstatudol. Nododd llawer o'r ymatebwyr i'r ymgynghoriad fanteision posibl dyletswydd statudol ar reoleiddwyr, ond cydnabu rhai y byddai cyflwyno'r gyfundrefn ar sail anstatudol yn y lle cyntaf yn caniatáu ar gyfer hyblygrwydd pwysig. Rydym o'r farn bod dull anstatudol yn cynnig hyblygrwydd hollbwysig ar hyn o bryd – yn enwedig tra byddwn yn dal i bennu ein dull gweithredu – ond byddwn yn parhau i adolygu hyn. Caiff ein penderfyniad ei lywio'n rhannol gan ein hadolygiad o'r cynlluniau a gyhoeddir gan y rheoleiddwyr, fel y nodir uchod; ein hadolygiad

---

<sup>25</sup> [AI Foundation Models: initial review](#), yr Awdurdod Cystadleuaeth a Marchnadoedd, 2023.

<sup>26</sup> [How do we ensure fairness in AI?](#), Swyddfa'r Comisiynydd Gwybodaeth, 2023.

<sup>27</sup> [Software and AI as a Medical Device Change Programme – Roadmap](#), yr Asiantaeth Rheoleiddio Meddyginiaethau a Chynhyrchion Gofal Iechyd, diweddarwyd 2023 [2021].

<sup>28</sup> Mae'r llywodraeth wedi ysgrifennu at y Swyddfa Gyfathrebiadau (Ofcom); Swyddfa'r Comisiynydd Gwybodaeth; yr Awdurdod Ymddygiad Ariannol; yr Awdurdod Cystadleuaeth a Marchnadoedd; y Comisiwn Cydraddoldeb a Hawliau Dynol; yr Asiantaeth Rheoleiddio Meddyginiaethau a Chynhyrchion Gofal Iechyd; y Swyddfa Safonau mewn Addysg, Gwasanaethau Plant a Sgiliau (Ofsted); y Bwrdd Gwasanaethau Cyfreithiol; y Swyddfa Rheoleiddio Niwclear; y Swyddfa Rheoleiddio Cymwysterau ac Arholiadau; yr Awdurdod Gweithredol Iechyd a Diogelwch (HSE); Banc Lloegr; a Swyddfa Marchnadoedd Nwy a Thrydan (Ofgem). Gofynnwyd hefyd i'r Swyddfa ar gyfer Diogelwch a Safonau Cynhyrchion, sy'n rhan o'r Adran Busnes a Masnach, gyflwyno diweddariad. Y rheoleiddwyr fydd yn y sefyllfa orau i bennu ffurf a sylwedd eu diweddariad ac rydym yn annog pob rheoleiddiwr sy'n ystyried bod deallusrwydd artiffisial yn berthnasol i'w waith i gyhoeddi ei ddull gweithredu. Wrth i ni barhau i roi'r fframwaith ar waith ac asesu parodrwydd rheoleiddwyr, efallai y bydd y ffordd rydym wedi blaenoriaethu rheoleiddwyr yn newid i adlewyrchu ffactorau sy'n datblygu megis ein dadansoddiad o risg. Byddwn hefyd yn gweithio gyda rheoleiddwyr eraill ac yn eu hannog i gyhoeddi cynlluniau gweithredu i sicrhau mwy o dryloywder yn yr ecosystem ehangach.

o bwerau rheoleiddwyr, fel y nodir isod; ac yn unol â'r dull gweithredu ehangach o ymdrin â deddfwriaeth ym maes deallusrwydd artiffisial, megis cyflwyno mesurau cyfrwymol wedi'u targedu (gweler adran 5.2).

### 5.1.2 Cefnogi galluedd a threfniadau cydgysylltu rheoleiddwyr

17. Mae'r newidiadau systemig sy'n cael eu sbarduno gan ddeallusrwydd artiffisial yn gofyn am ymateb system gyfan – ni all ein rheoleiddwyr unigol ymdrin â'r cyfleoedd a'r risgiau a gyflwynir gan dechnolegau deallusrwydd artiffisial o fewn eu cylchoedd gwaith yn llwyddiannus drwy weithredu ar eu pen eu hunain. Yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, cynigiwyd swyddogaeth ganolog newydd, i'w sefydlu yn y llywodraeth, er mwyn monitro ac asesu risgiau dros yr economi gyfan a chefnogi cydgysylltu ac eglurder rhwng rheoleiddwyr.

18. Croesawyd y cynnig i sefydlu swyddogaeth ganolog yn gyffredinol gan randdeiliaid, a nododd ei bod yn hollbwysig i gyflwyno'r fframwaith rheoleiddio deallusrwydd artiffisial yn effeithiol. Pwysleisiodd llawer, heb swyddogaeth o'r fath, fod risg o orgyffwrdd rheoleiddiol, bylchau a chydgyssylltu gwael wrth i nifer o reoleiddwyr ystyried effaith deallusrwydd artiffisial yn eu priod feysydd.

19. Rydym eisoes wedi dechrau sefydlu'r swyddogaeth hon mewn nifer o ffyrdd:

- i. **Asesu risg:** Rydym wedi recriwtio tîm amlddisgyblaethol newydd i ymgymryd â gwaith monitro risgiau traws-sector yn yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, sy'n dwyn ynghyd arbenigwyr ym maes risg, rheoleiddio, a deallusrwydd artiffisial, sydd â chefnidir mewn gwyddor data, peirianeg, economeg, a'r gyfraith. Bydd y tîm yn mynd ati i archwilio risgiau deallusrwydd artiffisial trawsbynciol yn barhaus, gan gynnwys gwerthuso effeithiolrwydd ymyriadau gan y llywodraeth a rheoleiddwyr. Yn 2024, byddwn yn lansio ymgynghoriad wedi'i dargedu ar gofrestr risgiau deallusrwydd artiffisial draws-economi er mwyn sicrhau ei bod yn cofnodi'r holl risgiau mewn ffordd gynhwysfawr. Bydd yn cynnig un ffynhonnell o wirionedd am risgiau deallusrwydd artiffisial y gall rheoleiddwyr, adrannau'r llywodraeth, a grwpiau allanol ei defnyddio. Bydd hefyd yn cefnogi gwaith y llywodraeth ar nodi unrhyw risgiau sy'n disgyn ar draws neu rhwng cylchoedd gwaith rheoleiddwyr fel y gallwn nodi lle mae'r bylchau neu lle mae'r trefniadau rheoleiddio presennol yn aneffeithiol a blaenoriaethu camau pellach. Yn ogystal â'r gofrestr risgiau, rydym yn ystyried gwerth ychwanegol datblygu fframwaith rheoli risg, yn debyg i un a ddatblygwyd yn yr UD gan y Sefydliad Cenedlaethol dros Safonau a Thechnoleg (NIST).
- ii. **Galluedd rheoleiddwyr:** Mae rheoleiddio effeithiol yn dibynnu ar reoleiddwyr sy'n meddu ar y sgiliau, yr adnoddau, a'r arbenigedd priodol. Er bod rhai rheoleiddwyr wedi llwyddo i roi'r arbenigedd priodol yn ei le i ymdrin â deallusrwydd artiffisial, mae eraill yn llai parod. Rydym yn cyhoeddi £10 miliwn i reoleiddwyr ddatblygu'r galluedd a'r adnoddau sydd eu hangen arnynt i ymaddasu ac ymateb i ddeallusrwydd artiffisial. Rydym yn buddsoddi yn ein rheoleiddwyr presennol er mwyn diogelu eu galluedd at y dyfodol. Drwy'r cyllid hwn, gall rheoleiddwyr gydweithio i greu, addasu, a gwella adnoddau ymarferol i ymdrin â risgiau a chyfleoedd deallusrwydd artiffisial o fewn ac ar draws eu cylchoedd gwaith. Bydd yn galluogi rheoleiddwyr i ymgymryd ag ymchwil a datblygu i lunio dirnadaethau newydd y gellir gweithredu arnynt, a fydd yn gosod sylfeini y ffordd y byddant yn gweithredu am flynyddoedd i ddod. Byddwn yn gweithio'n agos gyda rheoleiddwyr yn ystod y misoedd nesaf i nodi'r cyfleoedd mwyaf addawol i ddefnyddio'r cyllid hwn. Mae hyn yn adeiladu ar y cyhoeddiad diweddar y bydd y llywodraeth yn ystyried

sut i helpu rheoleiddwyr ymhellach i ddatblygu'r sgiliau arbenigol sydd eu hangen i reoleiddio technolegau sy'n cael eu datblygu, gan gynnwys opsiynau ar gyfer mwy o hyblygrwydd ynglŷn â chyflog ac amodau.<sup>29</sup>

- iii. **Pwerau rheoleiddwyr:** Rydym yn cydnabod bod angen asesu pwerau a chylchoedd gwaith presennol rheoleiddwyr y DU er mwyn sicrhau eu bod wedi'u paratoi i ymdrin â risgiau a chyfleoedd deallusrwydd artiffisial yn eu priod feysydd a gweithredu'r egwyddorion mewn ffordd gyson a chynhwysfawr. Felly, byddwn yn gweithio gydag adrannau'r llywodraeth a rheoleiddwyr i ddadansoddi ac adolygu bylchau posibl yn y pwerau a'r cylchoedd gwaith rheoleiddio presennol.
- iv. **Cydgysylltu:** Yn ystod y misoedd nesaf, byddwn yn ffurfioli ein gweithgareddau cydgysylltu â rheoleiddwyr. Er mwyn cefnogi a thywys y gwaith hwn, byddwn yn sefydlu pwyllgor llywio a fydd yn cynnwys cynrychiolwyr y llywodraeth a rheoleiddwyr allweddol er mwyn helpu i gyfnewid gwybodaeth a chydgyssylltu ynglŷn â llywodraethu ym maes deallusrwydd artiffisial erbyn gwanwyn 2024. Rydym yn parhau i gefnogi gweithgarwch cydgysylltu rheoleiddiol yn fwy cyffredinol, gan gynnwys gweithio gyda chyrrff megis y Fforwm Cydweithredu Rheoleiddio Digidol. Heddiw, rydym wedi cyhoeddi canllawiau newydd i reoleiddwyr er mwyn eu helpu i ddehongli a chymhwyso ein hegwyddorion.
- v. **Ymchwil ac arloesedd:** Rydym yn gweithio'n agos gydag Ymchwil ac Arloesi yn y DU (UKRI) er mwyn sicrhau y gall buddsoddiadau ehangach y llywodraeth mewn ymchwil a datblygu ym maes deallusrwydd artiffisial gefnogi agenda diogelwch y llywodraeth. Mae hyn yn cynnwys ymrwymiad newydd gan UKRI i wella cysylltiadau rhwng rheoleiddwyr a'r sgiliau, yr arbenigedd, a'r gweithgareddau a gefnogir gan fuddsoddiadau UKRI mewn ymchwil i ddeallusrwydd artiffisial megis Responsible AI UK, hwb Systemau Awtonomaidd Dibynadwy, Canolfannau Deallusrwydd Artiffisial UKRI ar gyfer Hyfforddiant Doethurol, a Sefydliad Alan Turing. Bydd hyn yn sicrhau bod cryfder y DU mewn ymchwil i ddeallusrwydd artiffisial yn cael ei ddefnyddio i'r eithaf yn ein fframwaith rheoleiddio. Mae'r gwaith hwn yn adeiladu ar ein hymrwymiad blaenorol o £250 miliwn drwy Gronfa Cenadaethau Technoleg UKRI i sicrhau arweinyddiaeth fyd-eang y DU mewn technolegau critigol.<sup>30</sup> Heddiw, mae UKRI yn cyhoeddi y bydd £19 miliwn o'r Gronfa Cenadaethau Technoleg yn cefnogi Cam 2 o'r gystadleuaeth ar Gyflymu Deallusrwydd Artiffisial Dibynadwy, gan gefnogi 21 o brosiectau a gyflwynir drwy raglen BridgeAI Innovate UK, er mwyn cyflymu'r broses o fabwysiadu deallusrwydd artiffisial dibynadwy a chyfrifol a dysgu peirianyddol.
- vi. **Hwylustod cydymffurfiaeth:** Mae'n rhaid i'r gyfundrefn reoleiddio weithio i arloeswyr. Rydym yn helpu arloeswyr a busnesau i gael cynhyrchion newydd ar y farchnad mewn ffordd ddiogel ac effeithlon drwy ariannu gwasanaethau cynghori amlasiantaethol peilot a gyflwynir gan y Fforwm Cydweithredu Rheoleiddio Digidol.<sup>31</sup> Bydd o gymorth arbennig i helpu arloeswyr i ddeall y gofynion cyfreithiol a rheoleiddiol y bydd angen iddynt eu bodloni cyn lansio cynnyrch. Bydd y porth ar-lein ar gyfer Hwb Deallusrwydd Artiffisial a Digidol peilot y Fforwm Cydweithredu Rheoleiddio Digidol a'r cyfnod ymgeisio yn lansio yn y gwanwyn. Bydd dirnadaethau o'r hwb peilot yn llywio'r ffordd y byddwn yn cyflwyno ein dull rheoleiddio. Mae

---

<sup>29</sup> [Response to Professor Dame Angela McLean's Pro-Innovation Regulation of Technologies Review: Cross Cutting](#). Trysorylys EF, 2023.

<sup>30</sup> [£250m to secure the UK's world-leading position in technologies of tomorrow](#), UKRI, 2023.

<sup>31</sup> Mae aelodau'r Fforwm Cydweithredu Rheoleiddio Digidol yn cynnwys yr Awdurdod Cystadleuaeth a Marchnadoedd, Swyddfa'r Comisiynydd Gwybodaeth, yr Awdurdod Ymddygiad Ariannol, ac Ofcom.

rhagor o fanylion am y meini prawf cymhwysra i gael y cymorth a gynigir o dan yr hwb peilot wedi cael eu cyhoeddi gan y Fforwm Cydweithredu Rheoleiddio Digidol heddiw, ochr yn ochr â'r ymateb hwn i'r ymgynghoriad.

- vii. **Ymddiriedaeth y cyhoedd:** Rydym am i fusnesau, defnyddwyr, a'r cyhoedd ymddiried mewn technolegau deallusrwydd artiffisial. Byddwn yn ennyn ymddiriedaeth drwy barhau i gefnogi gwaith ar dechnegau sicrwydd a safonau technegol. Mae Hwb Safonau Deallusrwydd Artiffisial y DU, a lansiwyd yn 2022, yn rhoi adnoddau a chanllawiau ymarferol i fusnesau, sefydliadau, ac unigolion ar ddefnyddio safonau technegol digidol a chymryd rhan yn y gwaith o'u datblygu.<sup>32</sup> Yn 2023, cydweithiodd y Llywodraeth â techUK i lansio'r Portffolio o Dechnegau Sicrwydd Deallusrwydd Artiffisial a gyhoeddwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial.<sup>33</sup> Yn ystod gwanwyn 2024, byddwn yn cyhoeddi "Introduction to AI assurance" er mwyn hyrwyddo gwerth sicrwydd deallusrwydd artiffisial ymhellach a helpu busnesau a sefydliadau i feithrin dealltwriaeth o'r technegau ar gyfer systemau diogel a dibynadwy. Och'r yn och'r â hyn, rydym yn cynnal ymchwil reolaidd ymhlith y cyhoedd er mwyn sicrhau bod y ffordd y mae'r llywodraeth yn gweithredu ynglŷn â deallusrwydd artiffisial yn gyson â'n gwerthoedd ehangach.<sup>34</sup>
- viii. **Monitro a Gwerthuso:** Rydym yn datblygu cynllun monitro a gwerthuso a fydd yn ein galluogi i asesu effeithiolrwydd ein cyfundrefn yn barhaus wrth i dechnolegau deallusrwydd artiffisial newid. Byddwn yn cynnal ymgynghoriad wedi'i dargedu â nifer o randdeiliaid ar ein cynllun arfaethedig i asesu'r fframwaith rheoleiddio yn ystod gwanwyn 2024. Fel rhan o'r gwaith hwn, byddwn yn ceisio barn fanwl ar ein metrigau a'n ffynonellau data arfaethedig.

20. Dim ond o fewn ecosystem ehangach sy'n hyrwyddo'r diwydiant y bydd rheoleiddio deallusrwydd artiffisial yn gweithio. Yn 2023, ymrwymodd y llywodraeth i ddarparu dros £1.5 biliwn i adeiladu uwchgyfrifiaduron yn y sector cyhoeddus, gan gynnwys yr Adnodd Ymchwil Deallusrwydd Artiffisial a chyfrifiadur ecsa-raddfa. Rydym hefyd yn gweithio'n agos gyda'r sector preifat i gefnogi buddsoddiad, megis cyhoeddiad Microsoft am £2.5 biliwn ar gyfer canolfannau data cysylltiedig â deallusrwydd artiffisial ym mis Tachwedd 2023. Bydd y buddsoddiad o £80 miliwn mewn hybiau deallusrwydd artiffisial rydym yn ei gyhoeddi heddiw yn ei gwneud yn bosibl i ddeallusrwydd artiffisial esblygu a mynd i'r afael â phroblemau cymhleth mewn cymwysiaidau o driniaethau gofal iechyd i electroneg ynni effeithlon. Mae'r llywodraeth hefyd yn cynnal adolygiad ehangach o gadwyn gyflenwi deallusrwydd artiffisial y DU er mwyn sicrhau ein bod yn cadw ein mantais strategol fel arweinydd byd-eang ar y technolegau hyn.

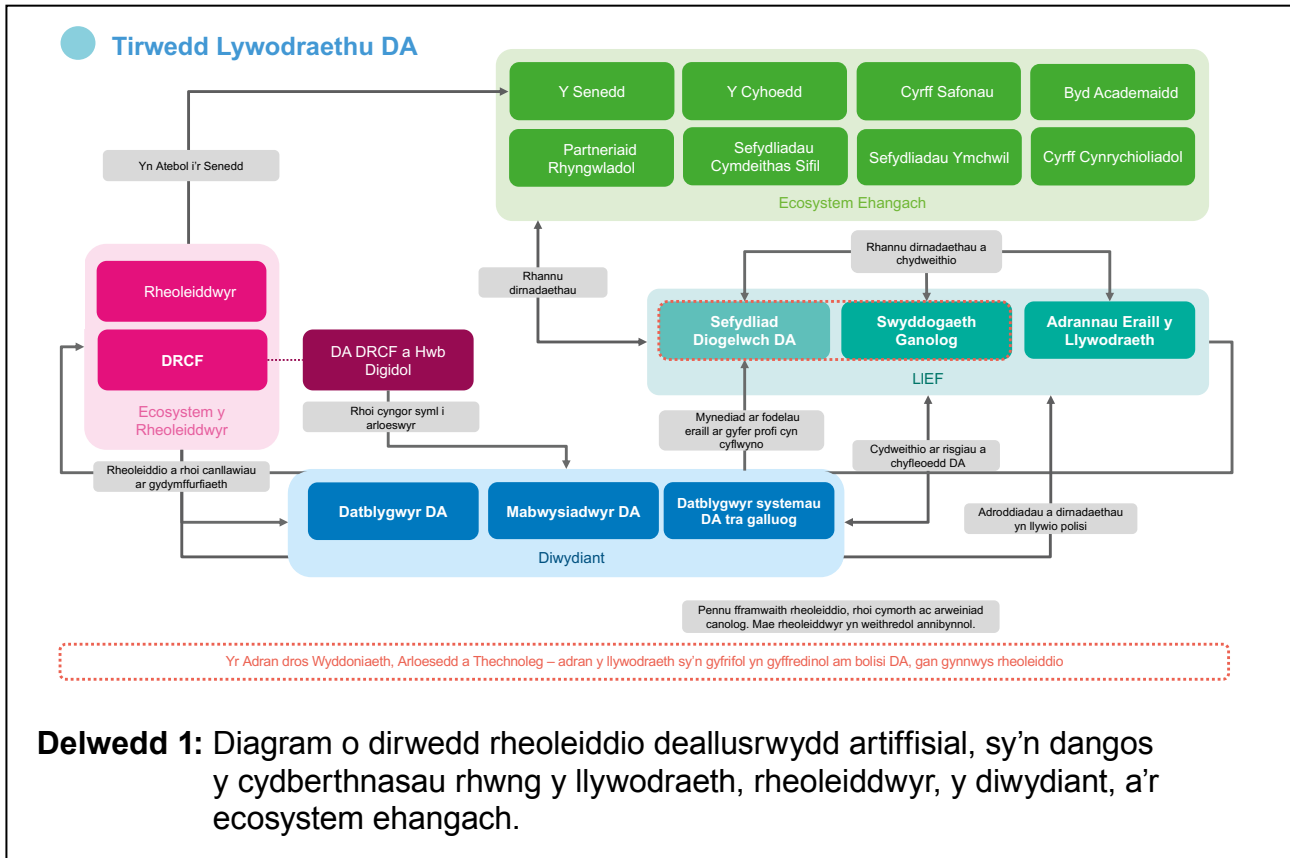
21. Yn olaf, er mwyn sbarduno pob rhan o'r llywodraeth i gymryd camau gweithredu cydgysylltiedig, rydym wedi pennu Gweinidogion deallusrwydd artiffisial arweiniol ym mhob adran i ddwyn ynghyd waith ar risgiau a chyfleoedd a ysgogir gan ddeallusrwydd artiffisial yn eu sectorau a goruchwylio'r gwaith o roi fframweithiau a chanllawiau ar waith ar gyfer defnydd y sector cyhoeddus o ddeallusrwydd artiffisial. Rydym hefyd wrthi'n sefydlu Grŵp Rhyngweinidogol newydd i sbarduno cydgysylltu effeithiol ym mhob rhan o'r llywodraeth ynglŷn â materion deallusrwydd artiffisial. At hynny, rydym yn cryfhau'r tîm sy'n gweithio ar ddeallusrwydd artiffisial yn yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg. Ym mis Chwefror 2023, roedd gennym dîm o ryw 20 o bobl yn gweithio ar faterion deallusrwydd artiffisial. Erbyn diwedd 2023, roedd y ffigur hwn wedi tyfu i fwy na 160 yn y Gyfarwyddiaeth Polisi Deallusrwydd Artiffisial newydd ei sefydlu a'r Sefydliad

<sup>32</sup> [The AI Standards Hub](#), Hwb Safonau Deallusrwydd Artiffisial, 2022.

<sup>33</sup> [Portfolio of AI Assurance Techniques](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>34</sup> [Public attitudes to data and AI: Tracker survey \(Wave 3\)](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023

Diogelwch Deallusrwydd Artiffisial, gyda'r bwriad o ehangu'r nifer hwn i fwy na 270 o bobl yn 2024. Er mwyn cydnabod y ffaith bod deallusrwydd artiffisial yn un o brif flaenoriaethau'r Ysgrifennydd Gwladol a'i fod wedi dod yn ganolog i waith ehangach yr adran a'r llywodraeth, ni fyddwn yn parhau i frandio Swyddfa Deallusrwydd Artiffisial ar wahân mwyach. Yn yr un modd, mae'r Ganolfan Moeseg Data ac Arloesi (CDEI) yn newid ei henw. Yr enw newydd fydd yr Uned Mabwysiadu Technoleg Cyfrifol i adlewyrchu ei chenhadaeth yn fwy cywir. Mae'r enw yn tynnu sylw at rôl y gyfarwyddiaeth i ddatblygu adnoddau a thechnegau sy'n ei gwneud yn bosibl i ddeallusrwydd artiffisial gael ei fabwysiadu mewn ffordd gyfrifol yn y sector preifat a'r sector cyhoeddus, er mwyn cefnogi cenhadaeth ganolog yr adran.



### 5.1.3 Mynd i'r afael â risgiau penodol

22. Mae tri chategori bras o risg sy'n gysylltiedig â deallusrwydd artiffisial: niweidiau cymdeithasol; risgiau camddefnydd; a risgiau awtonomi.<sup>35</sup> Amlinellwn isod enghreifftiau o'r ffordd y mae'r llywodraeth a'r rheoleiddwyr yn ymateb i risgiau penodol yn unol â'n hegwyddorion. Mae'r crynodeb hwn yn dangos yr amrywiaeth eang o waith sydd eisoes yn mynd rhagddo er mwyn sicrhau bod modd gwireddu manteision arloesi ym maes deallusrwydd artiffisial mewn ffordd ddiogel a chyfrifol. Ni fwriedir iddo fod yn gynhwysfawr na rhoi mwy o flaenoriaeth dros risgiau penodol nag eraill.

23. Yn ogystal â'r gwaith ar ymdrin â risgiau penodol a amlinellir isod, rydym heddiw yn cyhoeddi £2 filiwn o gyllid gan Gyngor Ymchwil y Celfyddydau a'r Dyniaethau er mwyn cefnogi ymchwil drosi a fydd yn helpu i ddiffinio deallusrwydd artiffisial cyfrifol mewn sectorau megis addysg, plismona, a'r diwydiannau creadigol. Bydd y prosiectau hyn, a fydd

<sup>35</sup> Rydym wedi categorio'r rhain yn flaenrol yn niweidiau cymdeithasol; risgiau camddefnydd; a cholli rheolaeth. Gweler: [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

yn rhan o waith BRAID (*Bridging Responsible AI Divides*) Cyngor Ymchwil y Celfyddydau a'r Dyniaethau,<sup>36</sup> yn llunio argymhellion i lywio gwaith yn y maes hwn yn y dyfodol a dangos sut mae'r DU o flaen y gad o ran ymgorffori deallusrwydd artiffisial mewn sectorau allweddol. Yn ogystal â'r prosiectau cwmpasu, mae Cyngor Ymchwil y Celfyddydau a'r Dyniaethau yn cadarnhau swm pellach o £7.6 miliwn i ariannu ail gam rhaglen BRAID, a fydd yn ymestyn gweithgareddau i 2027/28. Bydd y cam nesaf yn cynnwys carfan newydd o brosiectau arddangos ar raddfa fawr, cylchoedd pellach o Gymrodoriaethau BRAID, a darpariaethau sgiliau deallusrwydd artiffisial proffesiynol newydd, i'w datblygu ar y cyd â diwydiant a phartneriaid eraill.

## Niweidiau cymdeithasol

### Paratoi gweithwyr y DU ar gyfer economi wedi'i hwyluso drwy ddeallusrwydd artiffisial

24. Mae deallusrwydd artiffisial yn chwyldroi'r gweithle. Er y gellir creu swyddi newydd o ansawdd uwch drwy fabwysiadu'r technolegau hyn, gellir creu nifer o risgiau a'u gwneud yn fwy hefyd, megis gwyliadwriaeth yn y gweithle a gwahaniaethu wrth recriwtio, y mae'r llywodraeth a'r rheoleiddwyr eisoes yn gweithio i ymdrin â nhw. Rydym am fanteisio ar botensial twf deallusrwydd artiffisial ond ni ddylai hyn fod ar draul hawliau cyflogaeth na mesurau diogelu gweithwyr. Mae system gadarn y DU o ddeddfwriaeth a chamau gorfodi ynglŷn â mesurau diogelu cyflogaeth, gan gynnwys tribiwnlysoedd llafur arbenigol, yn gosod sylfaen gadarn i weithwyr. Er mwyn sicrhau bod y defnydd o ddeallusrwydd artiffisial ym maes adnoddau dynol a recriwtio yn ddiogel, yn gyfrifol ac yn deg, bydd yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg yn rhoi canllawiau wedi'u diweddarau yn ystod gwanwyn 2024. Mae system gadarn y DU o ddeddfwriaeth a chamau gorfodi ynglŷn â mesurau diogelu cyflogaeth, gan gynnwys tribiwnlysoedd llafur arbenigol, yn gosod sylfaen gadarn i weithwyr.

25. Ers 2018, rydym wedi ariannu pecyn gwerth £290 miliwn o fentrau sgiliau a thalent deallusrwydd artiffisial er mwyn sicrhau bod addysg ac ymwybyddiaeth o ddeallusrwydd artiffisial ar gael ledled y DU. Mae hyn yn cynnwys ariannu 24 o Ganolfannau Deallusrwydd Artiffisial ar gyfer Hyfforddiant Doethurol a fydd yn hyfforddi dros 1,500 o fyfyrwyr PhD. Rydym hefyd yn gweithio gydag Innovate UK a Sefydliad Alan Turing i ddatblygu canllawiau sy'n nodi'r sgiliau craidd sydd eu hangen ar bobl ym maes deallusrwydd artiffisial, o 'dinasyddion deallusrwydd artiffisial' i 'gweithwyr proffesiynol deallusrwydd artiffisial'. Ym mis Tachwedd 2023 gwnaethom gyhoeddi canllawiau drafft er mwyn cael sylwadau gan y cyhoedd a bwriadwn gyhoeddi fersiwn derfynol a fframwaith sgiliau llawn yn ystod gwanwyn 2024.<sup>37</sup>

26. Mae'n anodd rhagweld, ar hyn o bryd, sut yn union y bydd y farchnad lafur yn newid o ganlyniad i ddeallusrwydd artiffisial. Mae rhai sectorau yn pryderu y bydd deallusrwydd artiffisial yn disodli swyddi drwy awtomeiddio.<sup>38</sup> Mae'r Adran Addysg wedi cyhoeddi gwaith cychwynnol ar effaith deallusrwydd artiffisial ar swyddi, sectorau, cymwysterau, a llwybrau hyfforddiant yn y DU.<sup>39</sup> Gallwn fod yn hyderus y bydd angen sgiliau newydd arnom sy'n gysylltiedig â deallusrwydd artiffisial drwy gymwysterau cenedlaethol a darparu

---

<sup>36</sup> [Bridging Responsible AI Divides](#), BRAID UK, 2024.

<sup>37</sup> [AI Skills for Business Guidance: Feedback Consultation Call from The Alan Turing Institute](#), Innovate UK, 2023.

<sup>38</sup> Mae astudiaeth ddiweddar gan Sefydliad Dyfodol Gwaith yn dangos bod yr effaith net ar greu sgiliau a swyddi i gwmnïau yn y DU sydd wedi mabwysiadu technolegau deallusrwydd artiffisial a roboteg yn gadarnhaol. Fodd bynnag, mae'r effeithiau cadarnhaol hyn ar swyddi ac ansawdd swyddi yn gysylltiedig â lefelau parodrwydd mewn cwmni. Gweler: [What drives UK firms to adopt AI and robotics, and what are the consequences for jobs?](#), Sefydliad Dyfodol Gwaith, 2023.

<sup>39</sup> [The impact of AI on UK jobs and training](#), yr Adran Addysg, 2023.



hyfforddiant. Mae'r llywodraeth wedi buddsoddi £3.8 biliwn mewn addysg uwch ac addysg bellach yn ystod y tymor seneddol hwn i sicrhau bod y system sgiliau yn cael ei harwain gan gyflogwyr a'i bod yn ymateb i anghenion y dyfodol. Ochr ochr â phrentisaethau'r Adran Addysg<sup>40</sup> a Gwersylloedd Sgiliau,<sup>41</sup> bydd y diwygiadau i'r Hawl Dysgu Gydol Oes newydd<sup>42</sup> a'r Safon Brydeinig Uwch<sup>43</sup> yn rhoi addysg academaidd a thechnegol yn Lloegr ar sail gyfartal ac yn sicrhau bod ein system sgiliau ac addysg yn addas at y dyfodol.

## **Galluogi arloesi ym maes deallusrwydd artiffisial a diogelu eiddo deallusol**

27. Mae'r sector technoleg deallusrwydd artiffisial a'r sectorau creadigol, yn ogystal â'n cyfryngau, ar eu cryfaf pan fyddant yn gweithio gyda'i gilydd mewn partneriaeth. Mae'r llywodraeth hon yn ymrwymedig i gefnogi'r sectorau hyn er mwyn iddynt barhau i ffynnu a gallu cystadlu'n rhyngwladol. Mae'r Adran dros Dechnoleg Ddigidol, Diwylliant, y Cyfryngau a Chwaraeon yn gweithio'n agos gyda chyhoeddwyr, y diwydiant miwsig, a busnesau creadigol eraill i ddeall effaith deallusrwydd artiffisial ar y sectorau hyn, gyda'r nod o liniaru risgiau a manteisio ar gyfleoedd. Bydd cyllid sylweddol a nodwyd yng Ngweledigaeth Sector y Diwydiannau Creadigol<sup>44</sup> yn helpu i alluogi ymchwil a datblygu ac arloesi sy'n seiliedig ar ddeallusrwydd artiffisial yn y diwydiannau creadigol.

28. Mae gan y diwydiannau creadigol a sefydliadau yn y cyfryngau bryderon penodol ynglŷn â diogelu hawlfraint yn oes deallusrwydd artiffisial cynhyrchiol. Mae'r diwydiannau creadigol a deiliaid hawliau yn pryderu ynglŷn â'r defnydd o gynnwys wedi'i ddiogelu drwy hawlfraint ar raddfa fawr ar gyfer modelau deallusrwydd artiffisial hyfforddi ac maent wedi galw am sicrwydd y bydd eu gallu i gadw awtonomi a rheolaeth dros eu gwaith gwerthfawr yn cael ei ddiogelu. Ar yr un pryd, mae datblygwyr deallusrwydd artiffisial wedi pwysleisio bod angen iddynt allu cael mynediad hawdd at amrywiaeth eang o setiau data o ansawdd uchel i ddatblygu a hyfforddi systemau deallusrwydd artiffisial arloesol yn y DU.

29. Cynullodd y Swyddfa Eiddo Deallusol weithgor a oedd yn cynnwys deiliaid hawliau a datblygwyr deallusrwydd artiffisial ar ryngweithio rhwng hawlfraint a deallusrwydd artiffisial. Bu'r gweithgor yn fforwm gwerthfawr i randdeiliaid rannu eu safbwyntiau. Yn anffodus, mae bellach yn amlwg na all y gweithgor gytuno ar god gwirfoddol effeithiol.

30. Bydd gweinidogion yn yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg a'r Adran dros Dechnoleg Ddigidol, Diwylliant, y Cyfryngau a Chwaraeon bellach yn arwain cyfnod o ymgysylltu â'r sector deallusrwydd artiffisial a deiliaid hawliau, gyda'r nod o sicrhau dull gweithredu ymarferol ac effeithiol a fydd yn fodd i'r sector deallusrwydd artiffisial a'r sectorau creadigol dyfu gyda'i gilydd mewn partneriaeth. Mae'r llywodraeth yn ymrwymedig i dwf ein diwydiannau creadigol sy'n arwain y byd ac rydym yn cydnabod pwysigrwydd sicrhau bod datblygiadau deallusrwydd artiffisial yn cefnogi creadigrwydd dynol, arloesi, a'r gwaith o ddarparu gwybodaeth ddibynadwy, yn hytrach na'u tanseilio.

31. Bydd angen i'n dull gweithredu fod wedi'i seilio ar ymddiriedaeth a thryloywder rhwng partïon, gyda mwy o dryloywder ar ran datblygwyr deallusrwydd artiffisial mewn perthynas â mewnbynnau data, gyda rôl bwysig i briodoli allbynnau. Felly, bydd ein gwaith hefyd yn cynnwys ystyried systemau sy'n sicrhau mwy o dryloywder fel y gall deiliaid hawliau ddeall yn well a yw'r cynnwys a gynhyrchir ganddynt yn cael ei ddefnyddio fel mewnbyn mewn modelau deallusrwydd artiffisial. Mae'r llywodraeth am weithio'n agos gyda deiliaid hawliau

---

<sup>40</sup> Darperir prentisiaethau ar gyfer pobl 16 oed a throsodd nad ydynt mewn addysg amser llawn. Gweler: [Find an apprenticeship](#), yr Adran Addysg, dim dyddiad.

<sup>41</sup> Gwersylloedd Sgiliau i oedolion 19 oed ac yn hŷn. Gweler: [Find a skills bootcamp](#), yr Adran Addysg, 2024 [2022].

<sup>42</sup> [Lifelong Learning Entitlement overview](#), yr Adran Addysg, 2024.

<sup>43</sup> [A world-class education system: The Advanced British Standard](#), yr Adran Addysg, 2023.

<sup>44</sup> [Creative Industries Sector Vision](#), yr Adran dros Ddiwylliant, y Cyfryngau a Chwaraeon, 2023.

a datblygwyr deallusrwydd artiffisial i sicrhau hyn. O ran y gwaith hwn, bydd hefyd yn hollbwysig ymgysylltu'n agos â chymheiriaid rhyngwladol sydd hefyd yn mynd ati i ymdrin â'r materion hyn. Byddwn yn nodi cynigion pellach ar y ffordd ymlaen cyn hir.

### **Diogelu dinasyddion y DU rhag tuedd a gwahaniaethu sy'n gysylltiedig â deallusrwydd artiffisial**

32. Mae potensial i ddeallusrwydd artiffisial achosi i duedd a gwahaniaethu ymwreiddio ymhellach,<sup>45</sup> gan arwain o bosibl at ganlyniadau negyddol annheg i boblogaethau gwahanol mewn nifer o sectorau. Er enghraifft, gallai tuedd na chyfrifir amdani mewn proses gwneud penderfyniadau awtomataidd a hwylusir gan ddeallusrwydd artiffisial arwain at ganlyniadau gwahaniaethol yn erbyn nodweddion demograffig penodol mewn meysydd megis ceisiadau am gredyd<sup>46</sup> neu recriwtio.<sup>47</sup> Yn unol â'n hegwyddor tegwch, mae'r adran yn gweithio'n agos gyda'r Comisiwn Cydraddoldeb a Hawliau Dynol a Swyddfa'r Comisiynydd Gwybodaeth i ddatblygu atebion newydd i ymdrin â thuedd a gwahaniaethu mewn systemau deallusrwydd artiffisial.<sup>48</sup>

33. Mae'r ddau reoleiddiwr a chyrrff yn y sector cyhoeddus yn mynd i'r afael â thuedd a gwahaniaethu sy'n gysylltiedig â deallusrwydd artiffisial yn eu priod feysydd. Mae Swyddfa'r Comisiynydd Gwybodaeth wedi diweddarau canllawiau ynglŷn â sut mae ein cyfreithiau diogelu data cadarn yn gymwys i systemau deallusrwydd artiffisial sy'n prosesu data personol er mwyn cynnwys tegwch ac mae wedi parhau i ddwyn sefydliadau i gyfrif, er enghraifft drwy roi hysbysiadau gorfodi.<sup>49</sup> Cyhoeddodd Swyddfa Prif Gynghorydd Gwyddonol yr Heddlu Gyfamod ar Ddefnyddio Deallusrwydd Artiffisial ym maes Plismona<sup>50</sup> sydd wedi cael ei gymeradwyo gan Gyngor Prif Swyddogion yr Heddlu, y dylai pob un sy'n datblygu ac yn defnyddio'r dechnoleg yn y sector roi sylw priodol iddo.

### **Diwygio cyfraith diogelu data i gefnogi arloesedd a phreifatrwydd**

34. Data yw sylfaen modelu, hyfforddi, a datblygu systemau deallusrwydd artiffisial. Ond mae'n hollbwysig parchu hawliau unigolion perthnasol a dylid cydymffurfio ag egwyddorion diogelu data wrth brosesu data personol mewn systemau deallusrwydd artiffisial. Mae Swyddfa'r Comisiynydd Gwybodaeth wedi dangos sut y gall ddefnyddio cyfraith diogelu data i ddwyn sefydliadau i gyfrif drwy gamau rheoleiddio a chyfathrebu cyhoeddus lle mae systemau deallusrwydd artiffisial yn prosesu data personol. Bydd fframwaith diogelu data'r DU, sy'n cael ei ddiwygio drwy'r Bil Diogelu Data a Gwybodaeth Ddigidol, yn ategu ein dull cymesur o reoleiddio deallusrwydd artiffisial sydd o blaid arloesi ac sy'n seiliedig ar gyddestun.

35. Mae'r rheolau cyfredol ynglŷn â gwneud penderfyniadau awtomataidd yn ddryslyd ac yn gymhleth, sy'n tanseilio hyder i ddatblygu a defnyddio technolegau arloesol. Bydd y Bil Diogelu Data a Gwybodaeth Ddigidol yn ehangu ar y seiliau cyfreithlon dros wneud penderfyniadau cwbl awtomataidd sy'n cael effeithiau sylweddol ar unigolion ac yn rhoi hwb i hyder sefydliadau sy'n ystyried defnyddio'r technolegau mewn ffordd gyfrifol. Bydd yn parhau i sicrhau bod hawliau testun y data yn cael eu diogelu a bydd mesurau diogelu ar waith. Er enghraifft, bydd testunau data yn cael gwybodaeth am benderfyniadau o'r fath,

<sup>45</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>46</sup> [Algorithmic discrimination in the credit domain: what do we know about it?](#), Ana Cristina Bicharra Garcia et al., 2023.

<sup>47</sup> [Ethics and discrimination in artificial intelligence-enabled recruitment practices](#), Zhisheng Chen, 2023.

<sup>48</sup> [Fairness Innovation Challenge](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg; Innovate UK, 2023.

<sup>49</sup> [Guidance on AI and data protection](#), Swyddfa'r Comisiynydd Gwybodaeth, 2023.

<sup>50</sup> [Covenant for Using Artificial Intelligence \(AI\) in Policing](#), Cyngor Prif Swyddogion yr Heddlu, dim dyddiad.

yn cael cyfle i gyflwyno sylwadau, a byddant yn gallu gofyn am ymyriad dynol neu herio'r penderfyniad. Bydd hyn yn cefnogi arloesedd ac yn lleihau beichiau ar bobl a busnesau, ac yn parhau i ddiogelu data yn unol â safonau diogelu data uchel y DU ar yr un pryd.

### **Sicrhau bod cynnwys ar-lein a gynhyrchir drwy ddeallusrwydd artiffisial yn ddibynadwy ac yn ddiogel**

36. Mae'r llywodraeth yn ymrwymedig i sicrhau bod pobl yn cael mynediad at wybodaeth gywir ac mae'n cefnogi'r holl ymdrechion i hyrwyddo ffynonellau dilysadwy er mwyn mynd i'r afael â lledaenu gwybodaeth anwir neu gamarweiniol. Mae technolegau deallusrwydd artiffisial yn fwyfwy yn gallu cynnig ffyrdd rhad i bobl gynhyrchu cynnwys realistig a all bortreadu pobl a digwyddiadau ar gam. Yn yr un modd, gall deallusrwydd artiffisial gynyddu faint o gynnwys a gynhyrchir sy'n anfwriadol anwir, rhagfarnlyd, neu niweidiol.<sup>51</sup> Gall hyn ysgogi canfyddiadau negyddol ynglŷn ag ansawdd gwybodaeth ymhlith y cyhoedd a lleihau ymddiriedaeth gyffredinol mewn ffynonellau o wybodaeth.<sup>52</sup>

37. Rydym wedi cyhoeddi arferion datblygol i ddiogelu ymddiriedaeth mewn gwybodaeth ar-lein, gan gynnwys dyfrnodau a chronfeydd data allbynnau.<sup>53</sup> Byddwn yn lansio cais am dystiolaeth ynglŷn â'r risgiau sy'n gysylltiedig â deallusrwydd artiffisial i ymddiriedaeth mewn gwybodaeth cyn hir er mwyn meithrin ein dealltwriaeth o'r egin faes hwn o ddatblygu technolegol sy'n newid yn gyflym, gan gynnwys mesurau lliniaru posibl. Bydd hyn wedi'i anelu at ymchwilwyr, academyddion, a sefydliadau cymdeithas sifil sydd â'r arbenigedd perthnasol. Byddwn hefyd yn ystyried ymchwil i'r effeithiau ehangach a systemig ar yr ecosystem wybodaeth, a datrysiadau posibl. Rydym yn parhau i ymgysylltu â chyhoeddwyr newyddion a darlledwyr hefyd, fel sianeli hollbwysig o wybodaeth ddibynadwy a dilysadwy, ynglŷn â risgiau deallusrwydd artiffisial i newyddiaduraeth.

### **Sicrhau bod marchnadoedd digidol a lywir gan ddeallusrwydd artiffisial yn gystadleuol.**

38. Mae deallusrwydd artiffisial yn creu cyfleoedd enfawr i arloesi a fydd yn fuddiol i fusnesau a defnyddwyr drwy'r economi gyfan. Mae'r marchnadoedd ar gyfer y technolegau deallusrwydd artiffisial sylfaenol, megis modelau sylfaen, a'r cynhyrchion sy'n defnyddio deallusrwydd artiffisial mewn ffyrdd newydd ac arloesol, yn tyfu'n gyflym.

39. Pan fo'r marchnadoedd hyn yn gystadleuol byddant yn sbarduno arloesedd a gwell canlyniadau i fusnesau a defnyddwyr. Bydd cwmnïau llwyddiannus yn tyfu ac yn cynyddu eu cyfran o'r farchnad, a hynny'n gwbl briodol, ond bydd yn bwysig nad yw grym y farchnad yn ymwreiddio ymhellach ymhlith nifer bach o gwmnïau yn unig.

40. Bydd yr Awdurdod Cystadleuaeth a Marchnadoedd yn cymryd camau i sicrhau bod marchnadoedd deallusrwydd artiffisial yn gweithio'n dda i bawb. Ym mis Medi 2023, cyhoeddodd y rheoleiddiwr adolygiad cychwynnol o'r farchnad ar gyfer modelau sylfaen.<sup>54</sup> Canfu'r adroddiad, er y bydd llawer o fanteision i ddefnyddwyr o ganlyniad i ddeallusrwydd artiffisial, y gallai'r technolegau hyn ei gwneud yn bosibl i gwmnïau gipio grym yn y farchnad neu ymwreiddio ymhellach. Bydd y Bil Marchnadoedd Digidol, Cystadleuaeth a Defnyddwyr, sy'n mynd drwy Senedd y DU ar hyn o bryd, yn rhoi'r adnoddau ychwanegol i'r Awdurdod Cystadleuaeth a Marchnadoedd nodi unrhyw broblemau o ran cystadleuaeth mewn marchnadoedd deallusrwydd artiffisial a marchnadoedd digidol eraill yr effeithir arnynt gan ddatblygiadau diweddar ym maes deallusrwydd artiffisial ac ymdrin â nhw.

<sup>51</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>52</sup> [Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power](#), Katherine Ognyanova et al., 2020.

<sup>53</sup> [Emerging Processes for Frontier AI Safety](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>54</sup> [AI Foundation Models: initial review](#), yr Awdurdod Cystadleuaeth a Marchnadoedd, 2023.

## Sicrhau arferion gorau o ran deallusrwydd artiffisial yn y sector cyhoeddus

41. Mae deallusrwydd artiffisial yn cynnig cyfleoedd enfawr i drawsnewid cynhyrchiant yn y sector cyhoeddus. Mae'r DU eisoes yn arwain y ffordd, ac mae yn y trydydd safle ym Mynegai Paroddrwydd Llywodraethau ar gyfer Deallusrwydd Artiffisial.<sup>55</sup> Ym mis Tachwedd 2023, gwnaethom gyhoeddi ein bod yn treblu nifer y peirianwyr a'r datblygwyr deallusrwydd artiffisial technegol yn Swyddfa'r Cabinet i greu deorfa deallusrwydd artiffisial newydd i'r llywodraeth. Bydd yr arbenigwyr hyn yn dylunio ac yn gweithredu datrysiadau deallusrwydd artiffisial yn adrannau'r llywodraeth er mwyn sbarduno gwelliannau yn y ffordd y darperir gwasanaethau cyhoeddus. Gallai'r gwelliant posibl hwn mewn cynhyrchiant, er enghraifft, arbed hyd at 38 miliwn o oriau y flwyddyn a 750,000 o oriau bob wythnos i'r heddlu.<sup>56</sup>

42. Rydym yn achub ar y cyfleoedd sy'n codi o ganlyniad i ddeallusrwydd artiffisial i ddarparu gwell gwasanaethau cyhoeddus, gan gynnwys iechyd, addysg, a thrafnidiaeth. Er enghraifft, y llynedd, lansiodd yr Adran Iechyd a Gofal Cymdeithasol a'r GIG Gronfa Ddiagnostig Deallusrwydd Artiffisial gwerth £21m i gyflwyno'r technolegau hyn mewn meysydd allweddol lle ceir galw mawr megis pelydrau-x ar y frest a sganiau CT.<sup>57</sup> Mae'r Adran Addysg wedi bod yn ystyried sut i fanteisio i'r eithaf ar ddeallusrwydd artiffisial yn y sector addysg, gan gynnwys cyhoeddi papur polisi a chais am dystiolaeth ynglŷn â deallusrwydd artiffisial cynhyrchiol mewn addysg,<sup>58</sup> yn ogystal â rhedeg prosiect o ddigwyddiadau cyd-gyfrifiadura i ddeall achosion defnydd posibl ymhellach. Caiff canfyddiadau'r digwyddiadau cyd-gyfrifiadura mawr eu cyhoeddi yn ystod y gwanwyn eleni. Mae'r Adran Drafnidiaeth yn canolbwyntio ar y Bil Cerbydau Awtomeiddiedig newydd, sydd â'r nod o sicrhau bod y DU ar flaen y gad o ran rheoleiddio technoleg hunanyrru ac mewn sefyllfa gadarn i wireddu cyfran gwerth £42 biliwn amcangyfrifedig o'r farchnad hunanyrru fyd-eang. Mae'r Adran Drafnidiaeth hefyd yn bwriadu cyhoeddi ei Strategaeth Deallusrwydd Artiffisial Trafnidiaeth gyntaf yn 2024, er mwyn helpu'r adran a'r sector ehangach i ddeall y cyfleoedd a'r risgiau sy'n codi o ganlyniad i alluoedd deallusrwydd artiffisial newydd. Ochr yn ochr â hyn, mae'r adran yn parhau i ariannu busnesau bach a chanolig (BBaChau) arloesol drwy ei chynllun Grantiau Ymchwil ac Arloesi Trafnidiaeth i gefnogi'r genhedlaeth nesaf o adnoddau a chymwysiadau deallusrwydd artiffisial yn ogystal â threialu deallusrwydd artiffisial i helpu i nodi twyll yn ei phrosesau gwneud grantiau.

43. Mae Swyddfa'r Cabinet yn arwain y gwaith o osod y sylfeini angenrheidiol i ysgogi'r sector cyhoeddus cyfan i fabwysiadu deallusrwydd artiffisial, drwy wella seilwaith digidol a mynediad at setiau data, a datblygu safonau canolog. Mae'r llywodraeth hefyd yn defnyddio grym caffael y sector cyhoeddus i sbarduno datblygiadau deallusrwydd artiffisial arloesol cyfrifol a diogel. Mae'r Swyddfa Digidol a Data Ganolog wedi cyhoeddi canllawiau ar gaffael a defnyddio deallusrwydd artiffisial cynhyrchiol i Lywodraeth y DU.<sup>59</sup> Yn ddiweddarach eleni, bydd yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg yn lansio'r cynllun Hanfodion Rheoli Deallusrwydd Artiffisial, a fydd yn pennu safon arferion da ofynnol i gwmnïau sy'n gwerthu cynhyrchion a gwasanaethau deallusrwydd artiffisial. Byddwn yn ymgynghori ar gyflwyno'r safon hon fel gofyniad gorfodol ym maes caffael y sector cyhoeddus, gan ddefnyddio pŵer prynu i sbarduno ffyrdd cyfrifol o arloesi yn yr economi ehangach.

44. Mae hyn yn adeiladu ar y Safon Cofnodi Tryloywder Algorithmig, a sefydlwyd ffordd safonol i sefydliadau yn y sector cyhoeddus gyhoeddi gwybodaeth yn rhagweithiol ynglŷn â sut a pham y maent yn defnyddio dulliau algorithmig mewn prosesau gwneud penderfyniadau. Yn sgil treialu'r safon yn llwyddiannus, a chyhoeddi fersiwn

<sup>55</sup> [Government AI Readiness Index 2023](#), Oxford Insights, 2023.

<sup>56</sup> [Chancellor to cut admin workloads to free up frontline staff](#), Trysorlys EF; Y Swyddfa Gartref, 2023

<sup>57</sup> [£21 million to roll out artificial intelligence across the NHS](#), yr Adran Iechyd a Gofal Cymdeithasol, 2023.

<sup>58</sup> [Generative artificial intelligence in education call for evidence](#), yr Adran Addysg, 2023.

<sup>59</sup> [Generative AI Framework for HMG](#), Swyddfa'r Cabinet a'r Swyddfa Digidol a Data Ganolog, 2024.

drawslywodraethol gymeradwy y llynedd, byddwn yn ei gwneud yn ofynnol i bob un o adrannau'r llywodraeth ei defnyddio a bwriadwn ymestyn hyn i'r sector cyhoeddus yn fwy cyffredinol dros amser.

45. Er mwyn llywio'r defnydd diogel o ddeallusrwydd artiffisial ym mhob rhan o'r llywodraeth, y sector cyhoeddus, a thu hwnt, mae'r Ganolfan Seiberddiogelwch Genedlaethol wedi cyhoeddi nifer o ganllawiau ar yr ystyriaethau seiberddiogelwch sy'n codi wrth ddefnyddio a datblygu deallusrwydd artiffisial.<sup>60</sup>

## Risgiau camddefnydd

### Diogelu democratiaeth rhag ymyrraeth etholiadol

46. Mae'r llywodraeth yn ymrwymedig i gryfhau uniondeb etholiadau er mwyn sicrhau bod ein democratiaeth yn parhau'n ddiogel, yn fodern, yn dryloyw, ac yn deg. Mae gan ddeallusrwydd artiffisial y potensial i ymestyn cyrhaeddiad y rhai sy'n lledaenu twyllwbyodaeth ar-lein, i dargedu cynulleidfaoedd newydd yn fwy effeithiol, ac i gynhyrchu mathau newydd o gynnwys sy'n fwy anodd i'w canfod.<sup>61</sup> Mae ein Tasglu Amddiffyn Democratiaeth yn helpu i leihau bygythiad ymyrraeth estron yn ein democratiaeth drwy ddwyn ynghyd amrywiaeth eang o arbenigwyr o bob rhan o'r llywodraeth, y gymuned guddwbyodaeth, a diwydiant. Yn 2024, bydd y Tasglu yn ymgysylltu'n fwy helaeth â phartneriaid, gan gydweithio â'r llywodraethau datganoledig, yr heddlu, awdurdodau lleol, cwmnïau technoleg, a phartneriaid rhyngwladol.

47. Byddwn bob amser yn ymateb yn gadarn i unrhyw fygythiadau i ddemocratiaeth y DU. Cyflwynodd Deddf Etholiadau 2022 y gyfundrefn newydd o argraffnodau digidol, a fydd yn gwneud hysbysebion gwleidyddol digidol yn fwy tryloyw (gan gynnwys deunydd a gynhyrchir drwy ddeallusrwydd artiffisial), drwy ei gwneud yn ofynnol i'r rhai sy'n hyrwyddo deunydd ymgyrchu digidol cymwys a dargedir at etholwyr y DU gynnwys argraffnod sy'n cynnwys eu henw a'u cyfeiriad. Bydd hyn yn grymuso pleidleiswyr i wybod pwy sy'n hyrwyddo deunydd gwleidyddol ar-lein ac ar ran pwy. Gwnaeth Deddf Etholiadau 2022 hefyd ddiwygio'r drosedd o ddylanwad amhriodol. Bydd hyn yn diogelu pleidleiswyr yn well rhag dylanwadau amhriodol i bleidleisio mewn ffordd benodol, neu i beidio â phleidleisio o gwbl, ac mae'n cynnwys gweithgareddau sy'n twyllo person mewn perthynas â gweinyddiaeth etholiad (megis dyddiad etholiad neu leoliad gorsaf bleidleisio).

48. Bydd Deddf Diogelwch Ar-lein 2023 yn cwmpasu gweithgarwch penodol sydd wedi'i anelu at darfu ar etholiadau lle mae'n drosedd o fewn cwmpas y fframwaith rheoleiddio, megis cynnwys sy'n ysgogi trais yn erbyn ymgeiswyr etholiadol a ffigyrau cyhoeddus, a'r drosedd o ddylanwad amhriodol. Mae'r drosedd o ymyrraeth estron o Ddeddf Diogelwch Gwladol 2023 wedi cael ei hychwanegu at y Ddeddf Diogelwch Ar-lein fel "trosedd â blaenoriaeth", sy'n gosod cyfrifoldebau newydd ar ddarparwyr gwasanaethau ar-lein ac sy'n cwmpasu unrhyw ymgais gan y rhai sy'n gweithredu ar ran gwladwriaeth dramor i ymyrryd â'n hamgylchedd gwybodaeth a thanseilio ein prosesau democrataidd, gwleidyddol, a chyfreithiol (gan gynnwys etholiadau). Mae'r Ddeddf Diogelwch Ar-lein hefyd wedi diweddarau dyletswydd statudol Ofcom ynglŷn â llythrennedd cyfryngol, sy'n ei gwneud yn ofynnol i'r rheoleidiwr godi ymwybyddiaeth y cyhoedd o gamwbyodaeth a thwyllwbyodaeth ar-lein a'u helpu i'w gwrthsefyll.

49. Byddwn yn ystyried yr adnoddau sydd ar gael i ddilysu cynnwys sy'n ymwneud ag etholiadau. Gallai hyn gynnwys defnyddio dyfrnodau i roi hyder i bobl ynglŷn â'r cynnwys y maent yn ei weld. Nid dim ond y llywodraeth sydd angen gweithredu. Byddwn yn parhau i weithio gyda chwmnïau technoleg i sicrhau bod modd rhoi gwybod am gynnwys ffug

<sup>60</sup> [Artificial Intelligence](#), y Ganolfan Seiberddiogelwch Genedlaethol, dim dyddiad.

<sup>61</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

a'i ddileu'n gyflym. Gan adeiladu ar drafodaethau yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, rydym yn cydweithio â phartneriaid rhyngwladol a phartneriaid diwydiant i ymdrin â'r risg gyffredin o ymyrraeth mewn etholiadau.

### **Atal camddefnydd o dechnolegau deallusrwydd artiffisial**

50. Gall galluoedd deallusrwydd artiffisial gael eu defnyddio'n faleisus, er enghraifft i gyflawni seiberymosodiadau neu ddylunio arfau.<sup>62</sup> Gall datblygiadau ym maes deallusrwydd artiffisial gynyddu risgiau sy'n bodoli eisoes drwy ei gwneud yn bosibl i fygythwyr llai soffistigedig ymgymryd ag ymosodiadau mwy sylweddol ar raddfa fwy.<sup>63</sup> Rydym yn gweithio gyda diwydiant, academyddion, a phartneriaid rhyngwladol i ddod o hyd i fesurau lliniaru cymesur ac ymarferol i'r risgiau hyn. Bydd Strategaeth Diogelwch Biolegol 2023 ar ei newydd wedd yn sicrhau y bydd y DU, erbyn 2030, yn gallu gwrthsefyll rhychwant o risgiau biolegol ac yn arwain y byd o ran arloesi'n gyfrifol.<sup>64</sup> Fel y nodwyd yn y Weledigaeth Genedlaethol ar gyfer Bioleg Beirianyddol, mae'r llywodraeth wedi nodi sgrinio DNA synthetig fel blaenoriaeth polisi arloesi'n gyfrifol ar gyfer 2024.<sup>65</sup> Drwy flaenoriaethu hyn byddwn yn gallu parhau i fanteisio ar fuddiannau economaidd bioleg beirianyddol yn y DU a hefyd wella diogelwch y gadwyn gyflenwi ar yr un pryd.

51. Mae rhai o'r risgiau sy'n codi o ganlyniad i systemau deallusrwydd artiffisial yn dechrau dod i'r amlwg heddiw wrth i'r technolegau hyn gael eu camddefnyddio i gynyddu graddfa, cyflymder a llwyddiant troseddau. Fel y trafodir uchod, gall deallusrwydd artiffisial ei gwneud yn haws i ddefnyddwyr gynhyrchu cynnwys anwir neu gamarweiniol. Gall hyn gynnwys deunydd sy'n gyfystyr â throedd megis twyll, cam-drin plant yn rhywiol ar-lein, a chamddefnyddio delweddau personol. Mae'r llywodraeth eisoes wedi mynd ati i ymdrin â rhai o'r problemau hyn yn Neddf Diogelwch Ar-lein 2023. Gallai rhai technolegau deallusrwydd artiffisial gael eu camddefnyddio i gyflawni twyll hunaniaeth, megis cynhyrchu dogfennaeth ffug a ddefnyddir at ddibenion mewnffudo. Mae'r galluoedd hyn yn achosi risgiau posibl sy'n gysylltiedig â chael gafael ar gyllid cyhoeddus drwy dwyll.

52. Er mwyn ymdrin â'r defnydd troseddol o bosibl o ddeallusrwydd artiffisial, rydym yn adolygu i ba raddau y mae'r gyfraith droseddol bresennol yn darparu ar gyfer troseddau ac ymddygiad niweidiol a hwylusir gan ddeallusrwydd artiffisial. Gall deallusrwydd artiffisial hefyd achosi risgiau systemig i adnoddau'r heddlu, ymddiriedaeth sefydliadol, a'r broses dystiolaethol. Bydd y llywodraeth yn diwygio'r fframweithiau cyfreithiol presennol yn ôl y gofyn er mwyn diogelu cyfraith a threfn. Mae deallusrwydd artiffisial hefyd yn cynnig mwy o gyfleoedd ym maes gorfodi'r gyfraith i ddod yn fwy effeithlon o ran canfod ac atal troseddau. Fel y cyfryw, gall y technolegau hyn helpu i liniaru rhai o'r risgiau sy'n codi ynglŷn â throeddau a hwylusir gan ddeallusrwydd artiffisial. Er enghraifft, rydym yn buddsoddi mewn modelau deallusrwydd artiffisial sy'n galluogi'r heddlu i ganfod a chategoreiddio difrifoldeb delweddau o gam-drin plant yn fwy effeithiol. Rydym hefyd yn ystyried sut y gallai deallusrwydd artiffisial ei gwneud yn bosibl i swyddogion olygu darnau mawr o dystiolaeth ar ffurf testun yn gyflymach.

53. Er mwyn helpu sefydliadau i ddatblygu a defnyddio deallusrwydd artiffisial yn ddiogel, cyhoeddodd y Ganolfan Seiberddiogelwch Genedlaethol ganllawiau ar ddatblygu systemau deallusrwydd artiffisial diogel ym mis Tachwedd 2023. Mae'r llywodraeth bellach yn ystyried sut i adeiladu ar y rhain a chyhoeddiadau pwysig eraill drwy gyhoeddi cais am safbwyntiau yn ystod gwanwyn 2024 er mwyn cael mwy o gyfraniadau ar ein camau nesaf i sicrhau modelau deallusrwydd artiffisial, gan gynnwys Cod Ymarfer posibl ar seiberddiogelwch

---

<sup>62</sup> [Dual use of artificial-intelligence-powered drug discovery](#), Fabio Urbina et al., 2022.

<sup>63</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>64</sup> [UK Biological Security Strategy](#), Swyddfa'r Cabinet, 2023.

<sup>65</sup> [National vision for engineering biology](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

deallusrwydd artiffisial, yn seiliedig ar ganllawiau'r Ganolfan Seiberddiogelwch Genedlaethol. Mae cydweithio rhyngwladol yn y maes hwn yn hollbwysig os ydym am weld newid ystyrlon i ddiogelwch modelau deallusrwydd artiffisial, a byddwn yn ystyried ffyrdd o hyrwyddo cysondeb rhyngwladol, megis drwy safonau rhyngwladol.

54. Mae hyn yn adeiladu ar ein gwaith ar ddiogelu dyfeisiau personol a seilwaith allweddol. Bydd y gyfundrefn ddiogelwch yn y Ddeddf Diogelwch Cynhyrchion a Seilwaith Telathrebu, y bwriedir iddi ddod i rym yn 2024, yn ei gwneud yn ofynnol i weithgynhyrchwyr cynhyrchion cysylltadwy i ddefnyddwyr, megis seinyddion clyfar a hwylusir gan ddeallusrwydd artiffisial, gydymffurfio â gofynion diogelwch sylfaenol sy'n seiliedig ar yr egwyddor diogelu drwy ddylunio. Mae hyn yn golygu na fydd modd darparu unrhyw gynhyrchion cysylltadwy i ddefnyddwyr sydd o fewn cwmpas y gyfundrefn i gwsmeriaid yn y DU oni bai bod y gweithgynhyrchydd wedi rhoi mesurau diogelwch sylfaenol ar waith mewn perthynas â chaledwedd a meddalwedd y cynnyrch, a, lle y bo'n briodol, ddatrysiadau deallusrwydd artiffisial cysylltiedig. Y tu hwnt i hyn, mae'r Awdurdod Diogelwch Amddiffynnol Cenedlaethol yn cynnal ymchwil er mwyn deall sut y gall deallusrwydd artiffisial wella diogelwch ffisegol a diogelwch personél, a sut y bydd yn gwneud hynny. Mae'n cynghori amrywiaeth eang o sefydliadau, gan gynnwys cwmnïau seilwaith cenedlaethol allweddol, ynglŷn â sut i ymdrin â bygythiadau sy'n gysylltiedig â deallusrwydd artiffisial ac yn ymgyrchu i helpu i ddiogelu eiddo deallus gwerthfawr sy'n gysylltiedig â deallusrwydd artiffisial ar gyfer cwmnïau technoleg ddatblygol.

## Risgiau i awtonomi

55. Yn ein papur trafod ar alluoedd a risgiau deallusrwydd artiffisial ffin eithaf,<sup>66</sup> amlinellwyd risgiau posibl yn y dyfodol sy'n gysylltiedig ag awtonomi cynyddol systemau deallusrwydd artiffisial uwch. Mae rhai arbenigwyr yn pryderu, wrth i systemau deallusrwydd artiffisial ddod yn fwy galluog wrth ymgymryd ag amrywiaeth ehangach o dasgau, y bydd pobl yn dibynnu'n fwyfwy ar ddeallusrwydd artiffisial i wneud penderfyniadau pwysig. Mae rhai hefyd yn credu, yn y dyfodol, y bydd systemau deallusrwydd artiffisial â phŵer i weithredu yn gallu mynd ati'n weithredol i leihau rheolaeth ddynol a chynyddu eu dylanwad eu hunain. Mae ymchwil newydd i alluoedd datblygol deallusrwydd artiffisial â phŵer i weithredu yn dangos y bydd angen i ni ystyried mesurau newydd o bosibl i ymdrin â'r risgiau sy'n dod i'r amlwg wrth i dechnolegau deallusrwydd artiffisial sylfaenol sy'n sail i amrywiaeth o gymwysiadau barhau i ddatblygu.<sup>67</sup>

56. Yn adran 5.2, rydym yn nodi cynigion ar gyfer cyfrifoldebau newydd ar ddatblygwyr deallusrwydd artiffisial diben cyffredinol tra galluog yn y dyfodol. Tra bydd tebygolrwydd y risgiau i awtonomi yn cael ei drafod, credwn fod ein cynigion yn cyflwyno atebolrwydd, llywodraethu a goruchwyliaeth i'r datblygwyr hyn yn ogystal â phrofi a meincnodi systemau deallusrwydd artiffisial pwerus i ymdrin â'r risgiau hyn yn awr ac yn y dyfodol. Yn benodol, bydd profion a gynhelir gan y Sefydliad Diogelwch Deallusrwydd Artiffisial yn nodi systemau sydd â galluoedd peryglus o bosibl (gweler adrannau 5.2 a 5.3 am fwy o fanylion am rôl y Sefydliad). Mae profion eisoes wedi dechrau a bydd mwy ohonynt dros y misoedd dilynol. Mae'r camau cychwynnol hyn yn meithrin gallu technolegol y DU i asesu risgiau deallusrwydd artiffisial sy'n dod i'r amlwg ac ymateb iddynt, gan sicrhau ein cadernid o ran datblygiadau technolegol yn y dyfodol.

<sup>66</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>67</sup> [Practices for Governing Agentic AI Systems](#), Yonadav Shavit et al., 2023.

## 5.2. Ystyried yr achos dros gyfrifoldebau newydd i ddatblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog

57. Fel y nodir uchod, rydym yn gweld cynnydd cyflym o ran perfformiad systemau deallusrwydd artiffisial diben cyffredinol tra galluog. Rydym yn disgwyl i hyn barhau wrth i sefydliadau eu datblygu gyda mwy o gyfrifiannu, mwy o ddata, ac algorithmau mwy effeithlon. Nid yw datblygwyr bob amser yn gwybod beth fyddai galluoedd model cyn ei brofi.<sup>68</sup> Mae rhai cwmnïau wedi datgan yn gyhoeddus mai eu nod yw adeiladu systemau deallusrwydd artiffisial sy'n fwy galluog na bodau dynol wrth wneud nifer o dasgau.<sup>69</sup> Gyda galluoedd deallusrwydd artiffisial â phŵer i weithredu ar y gorwel, disgwyliwn newidiadau pellach a fydd yn trawsnewid ein cymdeithasau.<sup>70</sup>

58. Nododd Prif Weinidog y DU sut y byddai'r llywodraeth yn rheoli risg ar ffin eithaf datblygiadau deallusrwydd artiffisial ym mis Hydref 2023. Dywedodd: "My vision, and our ultimate goal, should be to work towards a more international approach to safety, where we collaborate with partners to ensure AI systems are safe before they are released."<sup>71</sup>

59. Nodwn isod sut mae'r DU wedi arwain y ffordd drwy ddull gweithredu technegol, gan sicrhau cytundebau gwirfoddol ynglŷn â diogelwch deallusrwydd artiffisial â gwledydd a chwmnïau allweddol. Bydd y Sefydliad Diogelwch Deallusrwydd Artiffisial newydd yn gweithio gyda'i bartneriaid i brofi systemau deallusrwydd artiffisial newydd mwy pwerus cyn ac ar ôl iddynt gael eu cyflwyno. Fel y nododd Prif Weinidog y DU, ni fyddwn yn brysio i reoleiddio a chyflwyno'r mesurau anghywir nad ydynt o bosibl yn sicrhau cydbwysedd digonol rhwng ymdrin â risgiau a chefnogi arloesedd.

60. Yn amlwg, os bydd galluoedd deallusrwydd artiffisial yn tyfu'n gyflymach ac yn gyflymach, ac os tybir bod mesurau gwirfoddol yn anghymesur â'r risg – fel y gallai ddigwydd yn ein barn ni – bydd gwledydd am weld rhai mesurau cyfrwymol i gadw'r cyhoedd yn ddiogel. Mae rhai gwledydd, megis yr Unol Daleithiau, yn dechrau archwilio hyn drwy ofynion adrodd gorfodol ar gyfer y systemau mwyaf pwerus. Rydym wedi gweld ffigyrau blaenllaw ym myd diwydiant, gwyddoniaeth, a chymdeithas sifil yn gwneud ymyriadau arwyddocaol, yn pwysleisio sut y dylai llywodraethau ystyried ymateb i gyflymder datblygiadau<sup>72</sup> ac rydym yn croesawu parhau i gydweithio agos â'r lleisiau arbenigol hyn.

61. Bydd y DU yn parhau i arwain y sgwrs ar lywodraethu deallusrwydd artiffisial yn effeithiol. Yn yr adran isod, nodwn rai o'r cwestiynau allweddol y bydd yn rhaid i wledydd fynd i'r afael â nhw wrth benderfynu ar y ffordd orau o reoli risgiau systemau deallusrwydd artiffisial diben cyffredinol tra galluog, megis sut i neilltuo atebolrwydd dros y gadwyn gyflenwi a negodi trefniadau rhyddhau agored ar gyfer y systemau mwyaf pwerus. Byddwn yn parhau i drafod y cwestiynau hyn â chymdeithas sifil, diwydiant, a phartneriaid rhyngwladol er mwyn paratoi at y dyfodol.

---

<sup>68</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>69</sup> [Race towards 'autonomous' AI agents grips Silicon Valley](#), Anna Tong a Jeffrey Dastin, 2023; [Introducing superalignment](#), Jan Leike a Ilya Sutskever (OpenAI), 2023; [AI could be one of humanity's most useful inventions](#), Google Deepmind, dim dyddiad.

<sup>70</sup> [Future Risks of Frontier AI](#), Swyddfa'r Llywodraeth dros Wyddoniaeth, 2023.

<sup>71</sup> [Prime Minister's speech on AI: 26 October 2023](#), Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.

<sup>72</sup> [Pause Giant AI Experiments: An Open Letter](#), Sefydliad Dyfodol Bywyd, 2023.



## **Blwch 2: Beth mae systemau deallusrwydd artiffisial diben cyffredinol tra galluog yn ei olygu?**

Yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, diffiniwyd “modelau sylfaen” fel math o fodel deallusrwydd artiffisial sy'n cael ei hyfforddi ar swm enfawr o ddata ac sy'n gallu cael ei addasu i'w ddefnyddio ar amrywiaeth eang o dasgau. Gall modelau sylfaen gael eu defnyddio fel sylfaen i adeiladu modelau deallusrwydd artiffisial mwy penodol.<sup>73</sup>

At ddibenion yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, gwnaeth y DU ddiffinio “deallusrwydd artiffisial ffin eithaf” fel modelau deallusrwydd artiffisial diben cyffredinol tra galluog sy'n gallu cyflawni amrywiaeth eang o dasgau ac sy'n meddu ar alluoedd cystal â'r modelau mwyaf datblygedig a welir heddiw neu sydd â galluoedd sy'n rhagori arnynt.

Heddiw, gall hyn gynnwys modelau sylfaen arloesol sy'n sail i gymwysiadau i ddefnyddwyr. Fodd bynnag, mae'n bwysig nodi y gallai systemau deallusrwydd artiffisial tra galluog, heddiw ac yn y dyfodol, fod yn seiliedig ar dechnoleg arall.

Yn yr ymateb hwn i'r ymgynghoriad, ffocws ein trafodaeth yw cyfrifoldebau yn y dyfodol ar gyfer datblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog. Ar hyn o bryd, mae datblygwyr y systemau hyn yn wynebu'r cyfrifoldebau cyfreithiol lleiaf eglur. Mae'r systemau wedi'u cwmpasu leiaf gan y trefniadau rheoleiddio presennol er eu bod yn achosi rhai o'r risgiau posibl mwyaf. Mae hyn yn golygu na fyddwn yn ymdrin â rhai o'r risgiau hynny'n effeithiol. Yn y dyfodol, efallai y bydd angen i'n dull rheoleiddio hefyd neilltuo cyfrifoldebau newydd i ddatblygwyr systemau cul tra galluog wrth i'r fframwaith barhau i ymaddasu i adlewyrchu datblygiadau technolegol newydd, risgiau gwahanol, neu ddadansoddiad pellach o atebolrwydd drwy gydol oes deallusrwydd artiffisial.

### **5.2.1. Yr heriau rheoleiddio sy'n gysylltiedig â deallusrwydd artiffisial diben cyffredinol tra galluog**

62. Amlinellodd y papur gwyn ar reoleiddio deallusrwydd artiffisial ddull rheoleiddio a oedd wedi'i gynllunio i addasu a datblygu ochr yn ochr â'r datblygiadau cyflym ym maes technoleg deallusrwydd artiffisial. Ar gyfer y mwyafrif helaeth o systemau deallusrwydd artiffisial, ein barn o hyd yw ei bod yn fwy effeithiol canolbwyntio ar y ffordd y defnyddir deallusrwydd artiffisial mewn cyd-destun penodol na rheoleiddio technolegau penodol, a hynny am y bydd lefel y risg yn cael ei phennu ar sail ble a sut y caiff deallusrwydd artiffisial ei ddefnyddio.

63. Fodd bynnag, gall rhai systemau deallusrwydd artiffisial tra galluog achosi risgiau sylweddol. Gall y risg gynyddu os mai system diben cyffredinol yw system dra galluog a bod modd ei defnyddio mewn amrywiaeth eang o gymwysiadau mewn sectorau gwahanol. Os bydd system deallusrwydd artiffisial diben cyffredinol yn achosi risg o niwed, gallai hyn olygu bod sawl sector neu gymhwysiad yn wynebu risg. Mae hyn yn golygu y gallai un nodwedd neu ddiffyg mewn un model arwain at niweidiau lluosog drwy'r economi gyfan. Er enghraifft, pe bai system deallusrwydd artiffisial yn cael ei defnyddio i ategu prosesau awtomataidd cymhleth ym maes gofal iechyd a recriwtio, ond mae allbynnau'r

<sup>73</sup> [AI regulation: a pro-innovation approach](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

model yn dangos tuedd mewn ffordd nad yw'n ddigon tryloyw neu maent yn cael effeithiau nad ydynt yn cael eu lliniaru'n ddigonol, gallai hyn arwain at arferion gwahaniaethol yn y gwasanaethau gwahanol hyn.

64. Mae systemau deallusrwydd artiffisial diben cyffredinol tra galluog yn herio dull o reoleiddio sy'n seiliedig ar gyd-destun gan na fydd rhai o'r risgiau y maent yn cyfrannu atynt o bosibl yn cael eu lliniaru'n effeithiol drwy'r trefniadau rheoleiddio presennol. Er enghraifft, efallai y bydd effaith draws-sector y systemau hyn yn ein hatal rhag gwneud digon i ymdrin â'r niweidiau hyn. Er y gall rhai rheoleiddwyr orfodi'r cyfreithiau presennol yn erbyn datblygwyr y systemau diben cyffredinol mwyaf galluog sydd o fewn eu cylchoedd gwaith cyfredol,<sup>74</sup> mae'r amrywiaeth eang o ddefnyddiau posibl yn golygu nad yw systemau diben cyffredinol yn cyd-fynd yn union â chylch gwaith unrhyw reoleiddiwr penodol ar hyn o bryd, a fydd o bosibl yn gadael risgiau heb unrhyw fesurau lliniaru effeithiol.<sup>75</sup>

65. Er bod rhai rheoleiddwyr yn dangos dulliau datblygedig o ymdrin â deallusrwydd artiffisial o fewn eu cylchoedd gwaith, mae'n bosibl na fydd llawer o'n fframweithiau cyfreithiol presennol na chylchoedd gwaith rheoleiddwyr yn lliniaru'n effeithiol y risgiau sy'n codi o ganlyniad i systemau deallusrwydd artiffisial diben cyffredinol tra galluog. Gall llawer o'r rheoleiddwyr yn y DU gael anhawster i orfodi'r rheolau presennol ar y rhai hynny sy'n dylunio, yn hyfforddi, ac yn datblygu'r systemau deallusrwydd artiffisial diben cyffredinol mwyaf pwerus. Yn yr un modd, nid yw'n amlwg bob amser sut y gellir cymhwyso'r rheolau presennol i ymdrin yn effeithiol â'r risgiau a all godi o ganlyniad i fodolau diben cyffredinol tra galluog. Caiff y rheolau a'r cyfreithiau presennol eu cymhwyso'n aml at lefel cyflwyno neu gymhwyso deallusrwydd artiffisial, ond mae'n bosibl na fydd y sefydliadau sy'n cyflwyno neu'n defnyddio'r systemau hyn mewn sefyllfa dda i nodi, asesu, neu lliniaru'r risgiau y gallant eu hachosi. Os felly, efallai y bydd cyfrifoldebau newydd ar ddatblygwyr modelau diben cyffredinol tra galluog yn ymdrin â risgiau'n fwy effeithiol.

66. Mae'n bosibl y bydd angen i'r gwaith sy'n mynd rhagddo gennym i ddadansoddi atebolrwydd cylch oes am ddeallusrwydd artiffisial a amlinellwyd yn y papur gwyn, yn y pen draw, ystyried rôl eraill sy'n rhan o'r gadwyn werth, megis darparwyr data neu gyfrifiadura cwmwl, er mwyn penderfynu sut y gellir dosbarthu cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial yn y ffordd decaf a mwyaf effeithiol. Bydd y dadansoddiad hwn hefyd yn ystyried sut y gallai'r ffordd anrhagweladwy y bydd galluoedd a risgiau yn y dyfodol yn datblygu amlygu bylchau eraill yn y dirwedd reoleiddio.

---

<sup>74</sup> Nodwn, er enghraifft, y camau gorfodi a gymerwyd gan Swyddfa'r Comisiynydd Gwybodaeth sydd wedi defnyddio cyfraith diogelu data i ddwyn sefydliadau sy'n defnyddio systemau deallusrwydd artiffisial sy'n prosesu data personol i gyfrif am dorri cyfraith diogelu data. Mae adolygiad cychwynnol yr Awdurdod Cystadleuaeth a Marchnadoedd o fodolau sylfaen yn nodi bod atebolrwydd am rwymedigaethau o dan gyfraith cystadleuaeth a chyfraith defnyddwyr yn gymwys gydol cylch oes deallusrwydd artiffisial i'r rhai sy'n ei ddatblygu ac sy'n ei gyflwyno fel ei gilydd. Gweler: [AI Foundation Models: initial review](#), yr Awdurdod Cystadleuaeth a Marchnadoedd, 2023. Yn yr un modd, mae Deddf Meddyginiaethau a Dyfeisiau Meddygol 2021 yn rhoi pwerau gorfodi i'r Asiantaeth Rheoleiddio Meddyginiaethau a Chynhyrchion Gofal Iechyd sy'n ddigonol i ddwyn gweithgynhyrchwyr dyfeisiau meddygol i gyfrif, gan gynnwys y pŵer i'w gwneud yn ofynnol iddynt dynnu dyfeisiau anniogel oddi ar y farchnad. Hefyd, gall camau gorfodi yn erbyn achosion difrifol o beidio â chydymffurfio, lle y bo'n briodol, arwain at erlyniad troseddol drwy'r llysoedd.

<sup>75</sup> Gall yr un model gael ei gyflwyno'n uniongyrchol gan y datblygwr a hefyd ei integreiddio mewn gwahanol systemau, cynhyrchion ac adnoddau dirifedi bron a fydd yn dod o dan gylch gwaith sawl rheoleiddiwr.

## **Astudiaeth achos 1: Atebolrwydd fel rhwystr i fabwysiadu deallusrwydd artiffisial yn y DU**

Mae "Count Your Pennies Ltd", sef cwmni cyfrifwyr dychmygol, yn prynu adnodd recriwtio deallusrwydd artiffisial "oddi ar y silff" a ddatblygwyd gan gwmni dychmygol yn y DU o'r enw "Quantum Talent Technologies". Mae'r adnodd yn llunio rhestrau byrion o ymgeiswyr yn awtomatig yn seiliedig ar eu ffurflenni cais.

Mae un ymgeisydd dychmygol, Ms Smith, yn holi pam y cafodd ei chais ei wrthod am swydd benodol o ystyried ei bod yn amlwg yn addas ar gyfer y rôl. Ar ôl cael ymateb anfodddhaol gan y rheolwr recriwtio, mae'n ffeilio hawliad am wahaniaethu. Drwy'r ymchwiliad, daw'n amlwg bod yr adnodd deallusrwydd artiffisial yn wahaniaethol. Cafodd ei adeiladu drwy ddefnyddio model sylfaen pwerus a ddatblygwyd gan gwmni y tu allan i'r DU a'i hyfforddi drwy ddefnyddio data cyflogaeth hanesyddol â thuedd.

Mae'n beth cyffredin i'r gyfraith neilltuo'r atebolrwydd i'r un olaf yn y gadwyn (yn yr achos hwn, "Count Your Pennies Ltd"). O dan amgylchiadau cyfyngedig, mae'n bosibl y bydd y gyfraith hefyd yn neilltuo atebolrwydd i'r sawl union uwchlaw yn y gadwn gyflenwi (yn yr achos hwn, "Quantum Talent Technologies").<sup>76</sup>

Er enghraifft, gall fod yn anodd i gyfraith cydraddoldeb – sy'n fframwaith statudol â'r nod o ddiogelu pobl rhag gwahaniaethu yn y gweithle ac mewn cymdeithas yn fwy cyffredinol o dan y gyfraith<sup>77</sup> – neilltuo atebolrwydd i unrhyw un heblaw'r cyflwynydd terfynol. Gallai hyn yn y pen draw arwain at ganlyniadau niweidiol (os nad oes cymhelliant gan y rhai sydd yn y sefyllfa orau i ymdrin â risgiau a niweidiau neu os nad ydynt yn cael eu dwyn i gyfrif) a thanseilio'r broses o fabwysiadu deallusrwydd artiffisial a gwanhau arloesedd yn economi'r DU. Byddwn yn parhau i ddadansoddi heriau megis y rhain fel rhan o'n gwaith polisi parhaus ar gyloch oes atebolrwydd am ddeallusrwydd artiffisial.

67. Er bod systemau deallusrwydd artiffisial cul tra galluog o fewn cwmnïau y fframwaith rheoleiddio ar gyfer deallusrwydd artiffisial, efallai y bydd y systemau hyn yn gofyn am set wahanol o ymyriadau os byddant yn datblygu galluoedd peryglus o bosibl. Mae systemau cul yn fwy tebygol o gael eu rheoleiddio'n effeithiol o fewn cylch gwaith rheoleiddiwr presennol na systemau diben cyffredinol. Byddwn yn parhau i gasglu tystiolaeth ynghylch a yw natur arbenigol systemau cul tra galluog yn gofyn am gamau gweithredu gwahanol i systemau diben cyffredinol.

<sup>76</sup> Efallai y bydd y gyfraith yn neilltuo atebolrwydd i "Quantum Talent Technologies" o dan y senario hwn os yw wedi cadarnhau perthynas "gweithredydd" yn unol â chyfraith cydraddoldeb neu os oedd rhwymedigaeth contract preifat arno i gydymffurfio â chyfraith cydraddoldeb. Mae'n bosibl y bydd y gyfraith hefyd yn priodoli atebolrwydd ar hyd y gadwyn gyflenwi mewn achos o esgeulustra os oes dyletswydd gofal sydd wedi cael ei thorri gan achosi niwed rhagweladwy. Fodd bynnag, dim ond i ryw rai sydd wedi'u lleoli yn y DU y bydd rhai cyfreithiau yn gymwys. O dan y senario hwn, byddai cyfraith diogelu data yn gymwys, a fyddai'n caniatáu i Swyddfa'r Comisiynydd Gwybodaeth gymryd camau gorfodi am unrhyw fethiant gan reolydd data perthnasol (megis "Count Your Pennies Ltd") i broses data personol mewn ffordd deg a chyfreithlon.

<sup>77</sup> [Deddf Cydraddoldeb 2010: guidance](#), Swyddfa Cydraddoldebau'r Llywodraeth a'r Comisiwn Cydraddoldeb a Hawliau Dynol, 2015 [2013].

## 5.2.2. Rôl mesurau gwirfoddol i greu dull rheoleiddio effeithiol ac wedi'i dargedu i gychwyn.

68. Rydym eisoes wedi dechrau gwneud y byd yn lle mwy diogel heddiw drwy sicrhau ymrwymadau gan gwmnïau deallusrwydd artiffisial blaenllaw ynglŷn â mesurau gwirfoddol. Gan adeiladu ar ymrwymadau gwirfoddol a sicrhawyd gan y Tŷ Gwyn, ysgrifennodd yr Ysgrifennydd Gwladol dros Wyddoniaeth, Arloesedd a Thechnoleg at saith cwmni deallusrwydd artiffisial ffin eithaf cyn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial yn gofyn iddynt gyhoeddi eu polisiau diogelwch. Cyhoeddodd pob un o'r saith cwmni eu polisiau cyn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, gan sicrhau mwy o dryloywder yn y gymuned deallusrwydd artiffisial a hyrwyddo ymarfer diogel yn y diwydiant.<sup>78</sup> Gwnaethom hefyd gyhoeddi adroddiad ar brosesau datblygol o ran diogelwch deallusrwydd artiffisial ffin eithaf er mwyn llywio'r broses o ddatblygu polisiau diogelwch yn y dyfodol (gweler Blwch 3).<sup>79</sup> Yn 2024, byddwn yn annog cwmnïau deallusrwydd artiffisial i ddatblygu eu polisiau o ran diogelwch deallusrwydd artiffisial a meithrin gallu ar raddfa gyfrifol.<sup>80</sup> Fel rhan o'r gwaith hwn, byddwn yn diweddarau ein canllaw ar brosesau datblygol erbyn diwedd y flwyddyn.

### **Blwch 3: Prosesau Datblygol ar gyfer Diogelwch Deallusrwydd Artiffisial Ffin Eithaf**

Cyn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, amlinellodd Llywodraeth y DU set o brosesau diogelwch datblygol er mwyn rhoi gwybodaeth i gwmnïau ynglŷn â sut y gallant sicrhau a chynnal diogelwch technolegau deallusrwydd artiffisial.

Mae'r ddogfen yn cwmpasu naw proses ddatblygol:

1. Meithrin Gallu ar Raddfa Gyfrifol – fframwaith ar gyfer rheoli risg wrth i sefydliadau feithrin gallu systemau deallusrwydd artiffisial ffin eithaf ar raddfa addas, sy'n galluogi cwmnïau i baratoi ar gyfer risgiau mwy peryglus, posibl yn y dyfodol o ran deallusrwydd artiffisial cyn iddynt godi.
2. Gwerthuso Modelau ac Ymarferion Tîm Coch – dulliau o asesu'r risgiau sy'n codi o ganlyniad i systemau deallusrwydd artiffisial a llywio gwell penderfyniadau ynglŷn â'u hyfforddi, eu sicrhau a'u cyflwyno.
3. Rhoi Gwybod am Fodelau a Rhannu Gwybodaeth – arferion sy'n ei gwneud yn haws i'r llywodraeth weld y deallusrwydd artiffisial ffin eithaf sy'n cael ei ddatblygu a'i gyflwyno ac sy'n ei gwneud yn bosibl i ddefnyddwyr wneud dewisiadau ar sail gwybodaeth ynghylch a ddylent ddefnyddio systemau deallusrwydd artiffisial a sut.
4. Rheolaethau Diogelwch, gan gynnwys Sicrhau Pwysau Modelau – mesurau megis seiberddiogelwch a rheolaethau diogelwch eraill sy'n sail i ddiogelwch systemau deallusrwydd artiffisial.
5. Strwythur Adrodd ar gyfer Gwendidau – proses sy'n galluogi pobl allanol i nodi problemau diogelwch mewn system deallusrwydd artiffisial.

<sup>78</sup> [Company Policies](#), yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, 2023.

<sup>79</sup> Emerging Processes for Frontier AI Safety, yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>80</sup> Fframwaith datblygol i reoli risgiau sy'n gysylltiedig â deallusrwydd artiffisial tra galluog a llywio prosesau gwneud penderfyniadau ynglŷn â datblygu a chyflwyno deallusrwydd artiffisial yw meithrin gallu ar raddfa gyfrifol. Gweler: [Responsible Capability Scaling](#) yn [Emerging Processes for Frontier AI Safety](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

6. Dynodwyr Deunydd a gynhrychir drwy Ddeallusrwydd Artiffisial – adnoddau i liniaru'r broses o greu a dosbarthu cynnwys twyllodrus a gynhrychir drwy ddeallusrwydd artiffisial drwy ddarparu gwybodaeth ynghylch a yw cynnwys wedi cael ei gynhyrchu neu ei addasu drwy ddeallusrwydd artiffisial.
7. Blaenoriaethu Ymchwil i Risgiau a achosir gan Ddeallusrwydd Artiffisial – prosesau ymchwil i nodi risgiau datblygol a achosir gan ddeallusrwydd artiffisial ffin eithaf ac ymdrin â nhw.
8. Atal a Monitro Camddefnydd o Fodelau – arferion i nodi ac atal camddefnydd bwriadol o systemau deallusrwydd artiffisial.
9. Rheolaethau ar Fewnbynnau Data ac Archwiliadau – mesurau i nodi a rheoli data hyfforddi sy'n debygol o gynyddu'r galluoedd peryglus sydd gan eu systemau deallusrwydd artiffisial ffin eithaf, a'r risgiau a achosir ganddynt.

Gwnaeth y ddogfen gydgrynhoi'r syniadau ynglŷn â diogelwch deallusrwydd artiffisial sy'n dod i'r amlwg ymhlith sefydliadau ymchwil a'r byd academiaidd, cwmnïau, a chymdeithas sifil, y gwnaeth Llywodraeth y DU gydweithio ac ymgysylltu â nhw drwy gydol y broses o'i datblygu. Prosiect parhaus yw diogelwch deallusrwydd artiffisial a bydd y prosesau a'r arferion yn parhau i ddatblygu drwy ymchwil a deialog rhwng llywodraethau a'r ecosystem deallusrwydd artiffisial ehangach. Mae'r ddogfen hon yn cynnig man cychwyn defnyddiol ar gyfer fframweithiau gweithredu yn y dyfodol yn y DU ac yn fyd-eang.

69. Ochr yn ochr â'r mesurau gwirfoddol hyn, yn yr Uwchgyhadledd ar Ddiogelwch Deallusrwydd Artiffisial, cytunodd llywodraethau a chwmnïau deallusrwydd artiffisial fod gan y naill barti a'r llall rôl hollbwysig i'w chwarae o ran profi'r genhedlaeth nesaf o fodelau deallusrwydd artiffisial er mwyn sicrhau diogelwch deallusrwydd artiffisial – cyn ac ar ôl i fodelau gael eu cyflwyno. Yn y DU, mae'r Sefydliad Diogelwch Deallusrwydd Artiffisial newydd ei sefydlu (gweler Blwch 4) yn arwain y gwaith hwn. Mae cwmnïau technoleg deallusrwydd artiffisial blaenllaw wedi addo rhoi mynediad blaenoriaethol i'r Sefydliad at eu systemau. Mae'r Sefydliad eisoes wedi dechrau cynnal profion, ac mae'n ymrwymedig i wneud hynny mewn partneriaeth â gwledydd eraill a'u sefydliadau diogelwch perthnasol. Byddwn yn rhoi diweddariad cyn hir ar y ffordd y mae'r Sefydliad Diogelwch Deallusrwydd yn cynnal gwerthusiadau. Bydd ein hasesiad o alluoedd a risgiau deallusrwydd artiffisial hefyd yn cael ei ategu gan Adroddiad Rhyngwladol newydd ar Wyddoniaeth Diogelwch Deallusrwydd Artiffisial,<sup>81</sup> o dan gadeiryddiaeth Joshua Bengio, sy'n arloesydd blaenllaw ym maes deallusrwydd artiffisial (gweler paragraff 87).

#### **Blwch 4: Y Sefydliad Diogelwch Deallusrwydd Artiffisial**

Ar hyn o bryd, mae datblygwyr deallusrwydd artiffisial ffin eithaf yn adeiladu systemau pwerus yn gyflymach na gallu'r llywodraeth a'r rheoleiddwyr i'w gwneud yn ddiogel. Fel y cyfryw, her gyntaf y llywodraeth yw gwybodaeth: nid ydym yn deall yn llawn yr hyn y gall y systemau mwyaf pwerus ei wneud ac mae angen dybryd i ni lenwi'r bwlch hwnnw. Hon fydd tasg y Sefydliad Diogelwch Deallusrwydd Artiffisial newydd. Bydd yn datblygu gwybodaeth y byd am ddiogelwch deallusrwydd artiffisial drwy archwilio, gwerthuso a phrofi systemau deallusrwydd artiffisial ffin eithaf yn ofalus. Hefyd, bydd

<sup>81</sup> [International expertise to drive International AI Safety Report](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2024.

yn ymchwilio i dechnegau newydd o ddeall a lliniaru risg deallusrwydd artiffisial, ac yn gwneud ymchwil sylfaenol ynglŷn â sut i gadw pobl yn ddiogel yn wyneb y cynnydd cyflym ac anrhagweladwy ym maes deallusrwydd artiffisial.

Bydd gwaith y Sefydliad Diogelwch Deallusrwydd Artiffisial yn hanfodol er mwyn llywio fframwaith rheoleiddio'r DU. Bydd yn rhoi dirnadaethau sylfaenol ynglŷn â'n cyfundrefn reoleiddio ac yn helpu i sicrhau bod y DU yn mabwysiadu dull cymesur o reoleiddio risgiau deallusrwydd artiffisial sy'n seiliedig ar risg. Bydd yn cyflawni tair swyddogaeth graidd i ddechrau:

- **Datblygu a chynnal gwerthusiadau o systemau deallusrwydd artiffisial uwch**, gyda'r nod o nodweddu galluoedd sy'n berthnasol i ddiogelwch, deall diogelwch systemau, ac asesu eu heffeithiau ar gymdeithas.
- **Sbarduno ymchwil sylfaenol i ddiogelwch deallusrwydd artiffisial**. Bydd ymchwil y Sefydliad yn cefnogi trefniadau llywodraethu deallusrwydd artiffisial yn y byrdymor a'r hirdymor. Bydd yn sicrhau bod fframwaith rheoleiddio iterus y DU ar gyfer deallusrwydd artiffisial yn cael ei lywio gan yr arbenigedd diweddaraf ac yn gosod y sylfeini ar gyfer trefniadau llywodraethu deallusrwydd artiffisial uwch rhyngwladol ac iddynt sylfaen dechnegol. Bydd prosiectau yn amrywio o ddatblygu adnoddau'n gyflym i lywio trefniadau llywodraethu, i ymchwil gychwynnol i ddiogelwch deallusrwydd artiffisial nad yw'r diwydiant o bosibl wedi rhoi sylw digonol iddo.
- **Hwyluso prosesau cyfnewid gwybodaeth**, gan gynnwys drwy sefydlu – ar sail wirfoddol ac yn ddarostyngedig i'r rheoliadau presennol ynglŷn â phreifatrwydd a data – sianeli clir i'r Sefydliad rannu gwybodaeth â chyfranogwyr cenedlaethol a rhyngwladol eraill, megis llunwyr polisi, partneriaid rhyngwladol, cwmnïau preifat, y byd academaidd, cymdeithas sifil, a'r cyhoedd yn fwy cyffredinol.

Ni fydd gwerthusiadau'r Sefydliad yn anelu at ddynodi unrhyw system deallusrwydd artiffisial benodol yn "ddiogel"; nid yw'n glir y gallai'r technegau sydd ar gael gyfiawnhau penderfyniad diffiniol o'r fath. Nid rheoleidiwr yw'r Sefydliad Diogelwch Deallusrwydd Artiffisial; ei rôl yw datblygu'r arbenigedd technegol i ddeall galluoedd a risgiau systemau deallusrwydd artiffisial, gan lywio camau gweithredu ehangach y llywodraeth. Serch hynny, rydym yn disgwyl iddo wneud cynnydd o ran gwerthuso systemau er mwyn galluogi llywodraethau a chwmnïau i wneud penderfyniadau gwell a gweithredu a gweithredu fel system rhybuddio cynnar i rai o'r risgiau sy'n peri'r pryder mwyaf. Os bydd y Sefydliad Diogelwch Deallusrwydd Artiffisial

yn nodi gallu peryglus o bosibl wrth werthuso systemau deallusrwydd artiffisial uwch, bydd yn ymdrin â'r risgiau, lle y bo'n briodol, drwy ymgysylltu â'r datblygwr ynglŷn â mesurau diogelwch lliniaru addas a chydweithio â phensaerniaeth rheoli risgiau deallusrwydd artiffisial a rheoleiddio sefydledig y llywodraeth.

Mae'r Sefydliad yn canolbwyntio ar y galluoedd deallusrwydd artiffisial presennol lefel uchaf ac unrhyw ddatblygiadau yn y dyfodol. Bydd yn ystyried systemau ffynhonnell agored yn ogystal â'r rhai a gyflwynir â gwahanol fathau o reolaethau mynediad.

70. Drwy'r camau gwirfoddol hyn bydd modd i ni brofi a dysgu beth sy'n gweithio er mwyn addasu ein dull rheoleiddio. Byddwn yn cryfhau ein dealltwriaeth dechnegol er mwyn creu consensws ehangach ynglŷn ag ymyriadau allweddol, er enghraifft, a ddylai fod amodau lle y byddai'n briodol atal datblygiad systemau penodol dros dro, fel y mae rhai wedi'i awgrymu.

71. Er bod y mesurau gwirfoddol yn ein helpu i wneud deallusrwydd artiffisial yn fwy diogel am y tro, mae'r gystadleuaeth ffyrnig rhwng cwmnïau i ryddhau systemau cynyddol alluog yn golygu y bydd angen i ni fod yn wyliadwrus iawn ynglŷn â chydymffurfiaeth ystyrion, atebolrwydd, a mesurau lliniaru risg effeithiol. Mae'n bosibl nad yw cymhellion masnachol bob amser yn gyson â budd y cyhoedd. Os bydd y farchnad yn datblygu yn y fath fodd nes bod mwy o gwmnïau yn adeiladu systemau tra galluog, bydd yn mynd yn llawer mwy anodd llywodraethu dulliau gwirfoddol.<sup>82</sup> Bydd hefyd yn gynyddol bwysig sicrhau bod y systemau atebolrwydd a'r fframweithiau llywodraethu corfforaethol priodol ar waith i gwmnïau sy'n adeiladu'r systemau mwyaf pwerus.

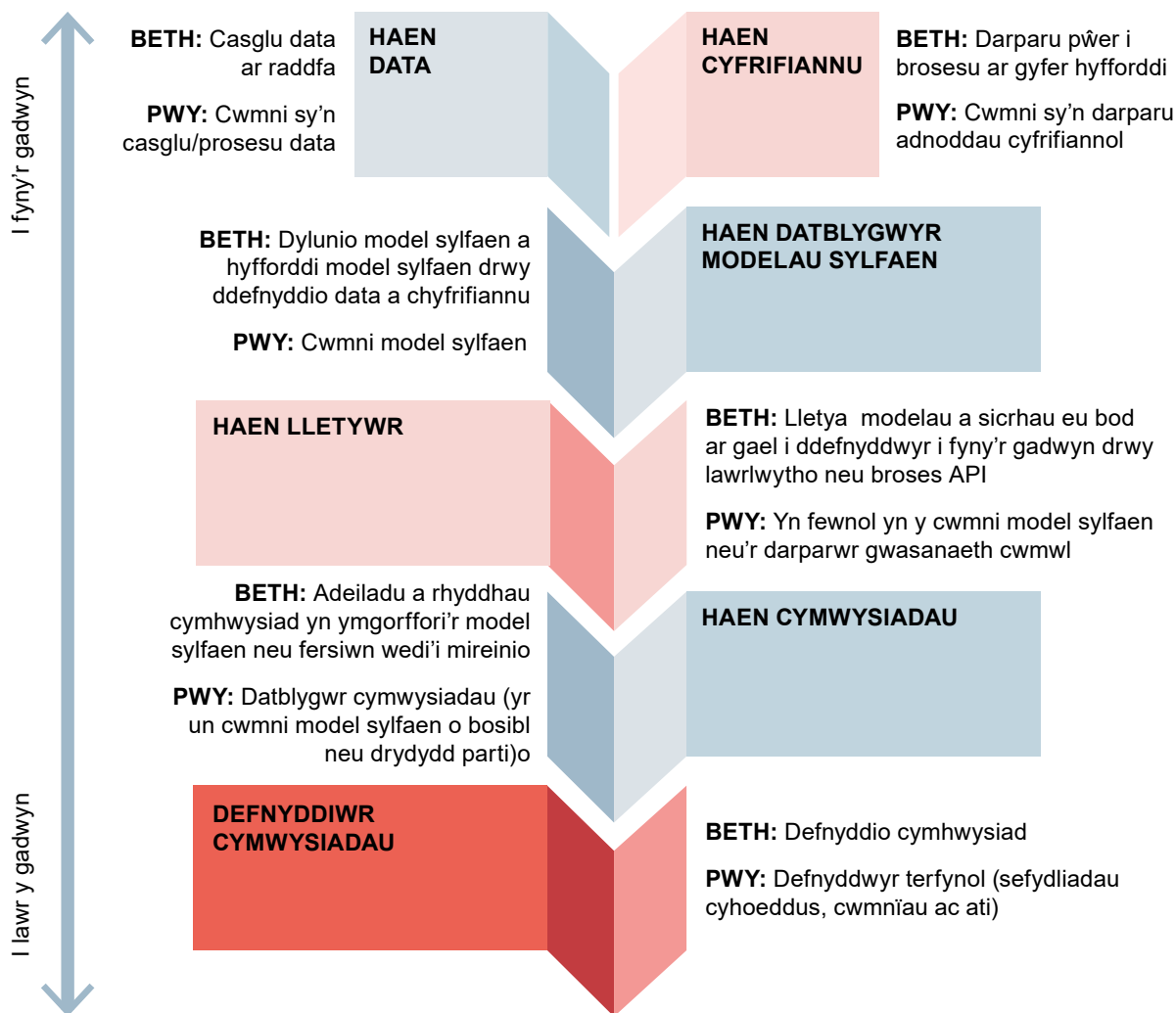
### 5.2.3. Yr achos dros fesurau cyfrwymol yn y dyfodol

72. Mae'r adran uchod yn tynnu sylw at y ffordd y gall dull gweithredu sy'n seiliedig ar gyd-destun golli risgiau sylweddol sy'n codi o ganlyniad i systemau diben cyffredinol tra galluog a gadael datblygwyr y systemau hynny yn anatebol. Er bod mesurau gwirfoddol yn ffordd ddefnyddiol o ymdrin â risgiau heddiw, disgwylwn y bydd pob awdurdodaeth, ymhen amser, yn awyddus i osod ymyriadau gorfodol wedi'u targedu ar y broses o ddylunio, datblygu a chyflwyno systemau o'r fath er mwyn sicrhau yr ymdrinnir â risgiau yn ddigonol.

---

<sup>82</sup> Er mwyn cefnogi gwaith cynllunio a datblygu polisiau'r llywodraeth, ac yn wyneb cryn ansicrwydd, mae Swyddfa'r Llywodraeth dros Wyddoniaeth wedi paratoi adroddiad rhagargoeli sy'n amlinellu senarios posibl a all godi yng nghyd-destun datblygu a lledaenu deallusrwydd artiffisial a'i effaith yn 2030. Gweler: [Future risks of frontier AI \(Annex A\)](#), Swyddfa'r Llywodraeth dros Wyddoniaeth, 2023. Caiff adroddiad llawn ar y senarios ei gyhoeddi cyn hir (ni fydd yr adroddiad hwn yn datgan polisi'r llywodraeth).

## Cadwyn gyflenwi modelau sylfaen



Noder: Model enghreifftiol yn unig yw hwn (efallai na fydd cwmni ar wahân neu unigol ar bob haen)

**Delwedd 2:** Diagram o gadwyn gyflenwi'r model sylfaen a ddaw o [Foundation Models Explainer](#) Sefydliad Ada Lovelace.

Er bod llawer o ffyrdd gwahanol o ddeall a disgrifio cylchoedd oes technolegau deallusrwydd artiffisial, sydd, yn aml, yn gymhleth, mae'r diagram hwn yn dangos y byddai ein mesurau arfaethedig yn y dyfodol yn amlwg yn cael eu targedu at nifer bach o gwmnïau sy'n gweithio ar haen datblygwyr modelau sylfaen sy'n adeiladu deallusrwydd artiffisial diben cyffredinol tra galluog.

73. Nid peth syml yw rhagweld pa systemau sy'n ddigon galluog i arwain at risg sylweddol. Yn unol â'n dull gweithredu cymesur, byddai unrhyw drefniadau rheoleiddio yn y dyfodol yn cael eu targedu at y nifer bach o ddatblygwyr sy'n adeiladu'r systemau diben cyffredinol mwyaf pwerus. Rydym yn cynnig y dylid gwneud hyn drwy bennu trothwyon dynamig sy'n gallu ymateb yn gyflym i ddatblygiadau ym maes deallusrwydd artiffisial. Mae ein



dadansoddiad rhagarweiniol yn dangos y gallai trothwyon cychwynnol fod yn seiliedig ar ragolygon o alluoedd drwy ddefnyddio cyfuniad o ddau fesur dirprwyol: cyfrifiannu (h.y. faint o gyfrifiannu a ddefnyddir i hyfforddi'r model) a meincnodi galluoedd (h.y. asesu galluoedd mewn meysydd risg penodol er mwyn nodi lle y bydd galluoedd lefel uchel yn arwain at risg uchel yn ein barn ni). Mae'r cyfuniad o'r ddau fesur dirprwyol hyn yn gallu rhagweld galluoedd deallusrwydd artiffisial yn eithaf da, am y tro, ond efallai y bydd angen ystod o drothwyon.

74. Byddai unrhyw rywmedigaethau newydd yn sicrhau bod datblygwyr y systemau sydd o fewn cwmpas yn cadw at yr egwyddorion a nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, gan gynnwys diogelwch, tryloywder, tegwch, ac atebolrwydd. Gallai hyn gynnwys mesurau tryloywder (erenghraifft, mesurau sy'n ymwneud â'r data a ddefnyddir i hyfforddi systemau); rheoli risg, atebolrwydd, a rhwymedigaethau sy'n gysylltiedig â llywodraethu corfforaethol; neu gamau gweithredu i ymdrin â niweidiau posibl, megis y rhai a achosir drwy gamddefnydd neu duedd annheg cyn neu ar ôl hyfforddiant.

75. Mae trefniadau rhyddhau agored ar gyfer deallusrwydd artiffisial wedi bod yn fuddiol ar y cyfan o ran arloesedd, tryloywder, ac atebolrwydd. Mae elfen o fod yn agored ym maes deallusrwydd artiffisial yn hollbwysig o ran datblygiadau gwyddonol, a bydd yn parhau felly, ac rydym yn cydnabod bod natur agored yn greiddiol i'n cymdeithas a'n diwylliant. Fodd bynnag, er ein bod yn ymrwymedig i amddiffyn gwerth bod yn agored, nodwn fod angen taro cydbwysedd wrth i ni geisio lliniaru risgiau posibl. Yn hyn o beth, rydym yn gweld consensws yn datblygu ynglŷn â'r angen i ystyried profion ar alluoedd cyn eu cyflwyno ac asesiadau risg ar gyfer y systemau deallusrwydd artiffisial mwyaf pwerus, gan gynnwys lle y gallai systemau gael eu rhyddhau'n agored. Gallai profion cyn cyflwyno lywio'r opsiynau cyflwyno sydd ar gael ar gyfer model a newid y camau atal risg y mae angen i sefydliadau eu cymryd cyn rhyddhau'r model. Gan gydnabod natur gymhleth y ddadl hon, rydym yn gweithio'n agos gyda'r gymuned ffynhonnell agored a datblygwyr deallusrwydd artiffisial i ddeall eu hanghenion. Drwy ymgysylltu â'r rhai sy'n datblygu ac yn defnyddio modelau deallusrwydd artiffisial sy'n rhai diben cyffredinol tra galluog, ac sydd â mynediad agored, byddwn yn gallu ystyried yr angen am opsiynau polisi cynnil ac wedi'u targedu sy'n lleihau unrhyw effeithiau negyddol ar weithgarwch ffynhonnell agored gwerthfawr i'r eithaf, a hefyd lliniaru risgiau ar yr un pryd.

76. Bydd yr heriau a ddaw yn sgil deallusrwydd artiffisial, yn y pen draw, yn gofyn am weithredu deddfwriaethol ym mhob gwlad unwaith y bydd y ddealltwriaeth o risg yn fwy aeddfed. Pe bai mesurau cyfrwymol yn cael eu cyflwyno'n rhy fuan, hyd yn oed rhai wedi'u targedu'n fanwl, gallent fethu ag ymdrin â risgiau'n effeithiol, mynd yn anghyfredol yn gyflym, neu lesteirio arloesedd ac atal pobl yn y DU rhag manteisio ar ddeallusrwydd artiffisial. Yn unol â'r dull hyblyg a nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, byddai'r llywodraeth yn ystyried cyflwyno mesurau cyfrwymol pe baem yn penderfynu nad oedd y mesurau lliniaru presennol yn ddigonol mwyach a phe baem wedi nodi ymyriadau a fyddai'n lliniaru risgiau mewn ffordd wedi'i thargedu. Fel yn achos unrhyw benderfyniad i ddeddfu, dim ond os nad oeddem yn ddigon hyderus y byddai mesurau gwirfoddol yn cael eu rhoi ar waith yn effeithiol gan bob parti perthnasol ac os oeddem wedi asesu na ellid lliniaru risgiau'n effeithiol drwy ddefnyddio'r pwerau cyfreithiol presennol y byddai'r llywodraeth yn ystyried cyflwyno deddfwriaeth. Yn olaf, cyn deddfu, byddai angen i'r llywodraeth fod yn hyderus y gallem orfodi mesurau mewn ffordd a fyddai'n lliniaru risg yn sylweddol heb amharu gormod ar arloesedd a chystadleuaeth.

77. Gwyddom fod mwy o waith i'w wneud i fireinio'r ffordd rydym yn rheoleiddio'r systemau deallusrwydd artiffisial mwyaf galluog a'r rhai sy'n eu dylunio, eu datblygu, a'u cyflwyno. Edrychwn ymlaen at ddatblygu ein cynigion drwy weithio'n agos gyda diwydiant, y byd academaidd, cymdeithas sifil, a'r cyhoedd yn fwy cyffredinol. Ym Mluch 5, isod, nodwn y cwestiynau allweddol a fydd yn llywio ein proses datblygu polisi.

### **Blwch 5: Cwestiynau allweddol ynglŷn â datblygu polisi ar reoleiddio systemau diben cyffredinol tra galluog yn y dyfodol**

Gan adeiladu ar y dystiolaeth a gawsom mewn ymateb i'n hymgyngoriad ar y papur gwyn ar reoleiddio deallusrwydd artiffisial ynglŷn ag atebolrwydd gydol cylch oes a modelau sylfaen, dros y misoedd nesaf byddwn yn gweithio'n agos gydag amrywiaeth o arbenigwyr a phartneriaid rhyngwladol i ystyried y cwestiynau isod. Byddwn yn cyhoeddi canfyddiadau'r gweithgarwch ymgysylltu hwn mewn cyfres o bapurau trafod arbenigol. Byddwn hefyd yn cyhoeddi'r iteriad nesaf o'n syniadau a'r camau rydym yn eu cymryd mewn perthynas â'r systemau deallusrwydd artiffisial mwyaf galluog.

- Pa risgiau penodol y dylid ymdrin â nhw drwy ymyriadau rheoleiddio yn y dyfodol sydd wedi'u targedu at systemau deallusrwydd artiffisial tra galluog? Sut rydym yn sicrhau bod y gyfundrefn yn gallu gwrthsefyll datblygiadau yn y dyfodol?
- Pryd y dylai'r llywodraeth a'r rheoleiddwyr ymyrryd? Pa systemau y dylem eu targedu? Sut beth fyddai trothwy cyfansawdd ar gyfer ymyriadau? A yw cyfrifiannu yn fesur dirprwyol defnyddiol am y tro, pe bai trothwyon yn parhau'n ddynamig? Beth am meincnodi galluoedd?
- Pa rwymedigaethau y dylid eu gosod ar ddatblygwyr? A ddylai'r rhwymedigaethau fod yn gysylltiedig â'n hegwyddorion rheoleiddio deallusrwydd artiffisial? Sut rydym yn sicrhau bod y rhwymedigaethau yn hyblyg ond yn glir? Ar ba gam y gallai fod yn angenrheidiol atal datblygiad model dros dro?
- Pa bwerau rheoleiddio newydd sydd eu hangen os o gwbl? Sut y byddai hyn yn gweithio ochr yn ochr â'r dirwedd reoleiddio bresennol?
- Sut y dylai unrhyw drefniadau rheoleiddio newydd gael eu gorfodi? Pa gyfrifoldebau cyfreithiol y dylid eu gosod ar ddatblygwyr systemau sydd o fewn cwmipas? A oes angen diweddarau'r fframweithiau atebolrwydd sifil neu droseddol?
- Sut y byddwn yn rhoi sicrwydd rheoleiddiol i sbarduno arloesi'n gyfrifol ym maes deallusrwydd artiffisial a hefyd gadw cyfundrefn hyblyg a all ddarparu ar gyfer datblygiadau technegol cyflym? Sut rydym yn osgoi creu rhwystrau sy'n atal cwmnïau rhag dod yn rhan o'r farchnad a datblygu ar raddfa addas?
- A ddylai galluoedd penodol ysgogi rheolaethau dros ryddhau agored? Beth fyddai'r goblygiadau negyddol? Sut y dylid pennu trothwyon? Pa reolaethau y dylid eu gosod?
- Beth yw rolau fframweithiau tryloywder ac atebolrwydd presennol? Sut y gellir annog neu sicrhau tryloywder cadarn ac atebolrwydd da er mwyn cefnogi datblygiad y systemau deallusrwydd artiffisial mwyaf galluog mewn ffordd gyfrifol?
- A ddylai datblygwyr systemau deallusrwydd artiffisial tra galluog fod yn ddarostyngedig i ofynion llywodraethau corfforaethol penodol? A all gofynion ar ddatblygwyr systemau deallusrwydd artiffisial tra galluog i ystyried a lliniaru risgiau i gymdeithas neu'r ddynolryw yn gyffredinol chwarae rôl?

- Beth yw'r berthynas rhwng mesurau newydd posibl ynglŷn â systemau deallusrwydd artiffisial tra galluog ac atebolrwydd cylch oes ehangach am ddeallusrwydd artiffisial? A oes unrhyw rai eraill yng nghadwyn werth deallusrwydd artiffisial y mae hefyd yn anodd i reoleiddwyr eu cyrraedd mewn ffordd sy'n llesteirio ein gallu i ymdrin â risgiau a chefnogi arloesedd mewn deallusrwydd artiffisial a'i fabwysiadu?

78. Fel y nodwn yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, ein bwriad yw y bydd ein fframwaith rheoleiddio yn gymwys i'r DU gyfan yn ddarostyngedig i esemptiadau a rhanddirymiadau ar gyfer gofynion gweithredu unigryw, megis amddiffyn a diogelwch gwladol. Fodd bynnag, rydym yn cydnabod bod deallusrwydd artiffisial yn cael ei ddefnyddio mewn amrywiaeth eang o sectorau, y mae rhai ohonynt wedi'u cadw'n ôl ac y mae eraill wedi'u datganoli. Wrth i'n polisi ddatblygu, ac wrth i ni ystyried cyflwyno gofynion cyfrwymol ar ddatblygwyr y systemau diben cyffredinol mwyaf galluog, byddwn yn parhau i asesu unrhyw effeithiau ar ddatganoli a'r angen am gyrhaeddiad alldiriogaethol.

79. Rydym yn ymrwymedig i ymgysylltu â'r swyddfeydd tiriogaethol a'r gweinyddiaethau datganoledig ynglŷn â llunio a chyflwyno'r fframwaith rheoleiddio, er mwyn i fusnesau a dinasyddion ledled y DU gael budd o'n dull rheoleiddio.

### 5.3. Gweithio gyda phartneriaid rhyngwladol i hyrwyddo cydweithrediad effeithiol ar lywodraethu deallusrwydd artiffisial

80. Nid oes unrhyw ffiniau i ddeallusrwydd artiffisial a bydd ei effaith yn dylanwadu ar gymdeithasau ac economïau ym mhedwar ban byd: Bydd deallusrwydd artiffisial a ddatblygir mewn un wlad yn effeithio'n gynyddol ar fywydau dinasyddion sy'n byw mewn gwledydd eraill. Felly, er mwyn llywodraethu deallusrwydd artiffisial yn effeithiol, bydd angen i gydweithrediad rhyngwladol fod yr un mor effeithiol, a bydd yn rhaid iddo adeiladu ar waith fforymau a mentrau amlochrog ac aml-randdeiliaid sy'n bodoli eisoes.

81. Mae'r DU yn arweinydd byd-eang sefydledig ym maes deallusrwydd artiffisial, gyda hanes o ysgogi trafodaethau rhyngwladol a chymryd camau clir a phenderfynol i greu cytundeb dwyochrog ac amlochrog. Mae ein ffocws hyd yma wedi bod ar gamau cydweithredol i helpu i ddatblygu deallusrwydd artiffisial yn unol â'r fframwaith sy'n seiliedig ar gyd-destun a'r egwyddorion a nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial.<sup>83</sup> Mae hyn yn golygu gweithio ochr yn ochr â gwahanol grwpiau o wledydd yn ôl yr angen a gweithredu mewn modd cymesur ac wedi'i dargedu. Ein nod o hyd yw gweithio gydag eraill i greu cymuned ryngwladol sy'n gallu gwireddu cyfleoedd deallusrwydd artiffisial ar raddfa byd-eang. Rydym yn hyrwyddo ein gwerthoedd ac yn cydweithio lle y bo'n addas er mwyn ymdrin â'r risgiau mwyaf dybryd sy'n gysylltiedig â deallusrwydd artiffisial ar hyn o bryd ac yn y dyfodol. Rydym yn cydbwyso diogelwch ac arloesedd yn ofalus, gan weithredu ochr yn ochr â phartneriaid i hyrwyddo proses ryngwladol o ddylunio, datblygu, cyflwyno a defnyddio'r systemau deallusrwydd artiffisial â'r potensial mwyaf.

82. Byddwn yn parhau i weithredu drwy bartneriaethau dwyochrog a mentrau amlochrog – gan gynnwys Uwchgynadleddau ar Ddiogelwch Deallusrwydd Artiffisial yn y dyfodol – i hyrwyddo deallusrwydd artiffisial diogel a dibynadwy, sydd wedi'i ategu gan drefniadau llywodraethu deallusrwydd artiffisial rhyngwladol effeithiol. Drwy'r cyfan byddwn yn

<sup>83</sup> [AI regulation: a pro-innovation approach](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

gweithredu mewn ffordd sy'n cynnwys nifer o randdeiliaid: Byddwn yn cydweithio â'n partneriaid rhyngwladol drwy weithio gyda chynrychiolwyr o ddiwydiant, y byd academaidd, cymdeithas sifil, a'r llywodraeth er mwyn sicrhau ein bod yn gallu sicrhau'r manteision eithriadol a gynigir gan y technolegau hyn.<sup>84</sup>

83. Gan weithio gyda'r rhwydweithiau hyn, byddwn yn datgloi'r cyfleoedd sy'n codi o ganlyniad i ddeallusrwydd artiffisial a hefyd ymdrin â risgiau posibl ar yr un pryd. Er mwyn cefnogi hyn, byddwn yn cynnal cydberthnasau agos â'n partneriaid rhyngwladol o ran yr holl faterion a nodir yn adran 5.1, yn ogystal â'n priod ddulliau gweithredu domestig sy'n datblygu.

84. Mae'n rhaid i ddulliau gweithredu domestig a rhyngwladol ddatblygu ar y cyd. Wrth ddatblygu ein dull ein hunain o reoleiddio deallusrwydd artiffisial, byddwn felly yn dylanwadu ar ddatblygiadau rhyngwladol ac yn ymateb iddynt. Byddwn yn parhau i ymgysylltu'n rhagweithiol yn y dirwedd ryngwladol er mwyn sicrhau'r lefel briodol o gydweithredu sydd ei hangen i lywodraethu deallusrwydd artiffisial yn effeithiol. Byddwn yn sicrhau lefelau priodol o gydlyniant â chyfundrefnau rheoleiddio eraill, yn hyrwyddo diogelwch, ac yn lleihau'r rhwystrau posibl i fasnach – gan sicrhau'r cyfleoedd mwyaf posibl i unigolion a busnesau ledled y DU a thu hwnt. Byddwn yn parhau i weithio gyda'n partneriaid rhyngwladol i sbarduno'r gwaith o ddatblygu a mabwysiadu adnoddau ar gyfer deallusrwydd artiffisial dibynadwy, megis technegau sicrwydd a safonau technegol bydeang, er mwyn hyrwyddo'r gallu i ryngweithredu ac osgoi trefniadau darniog.

85. Byddwn yn parhau i gydnabod natur hollbwysig diogelwch er mwyn ategu, ond nid disodli, pob agwedd arall ar gydweithio rhyngwladol ar ddeallusrwydd artiffisial. Fel y nododd Prif Weinidog y DU, Rishi Sunak “[our] vision, and our ultimate goal, should be to work towards a more international approach to safety”.<sup>85</sup> Fel y nodir uchod, cynhaliodd y DU yr Uwchgynhadledd gyntaf erioed ar Ddiogelwch Deallusrwydd Artiffisial ym mis Tachwedd 2023, gan sicrhau Datganiad Bletchley, sef cytundeb pwysig rhwng 29 o bartion, gan gynnwys 28 o wledydd o bedwar ban byd a'r Undeb Ewropeaidd.<sup>86</sup> Mae'r Datganiad yn meithrin dealltwriaeth gyffredin o'r cyfleoedd a'r risgiau sy'n codi o ganlyniad i ddeallusrwydd artiffisial a'r angen am gamau cydweithredol er mwyn sicrhau diogelwch y systemau deallusrwydd artiffisial mwyaf pwerus, yn awr ac yn y dyfodol. Cytunodd nifer o wledydd a chwmnïau sy'n datblygu deallusrwydd artiffisial ffin eithaf hefyd â phroffion ar y genhedlaeth nesaf o systemau a arweinir gan wladwriaethau, gan gynnwys drwy bartneriaethau â Sefydliadau Diogelwch Deallusrwydd Artiffisial newydd eu cyhoeddi (gweler Blwch 4 am ragor o fanylion).<sup>87</sup>

86. Nid oes unrhyw arwydd bod datblygiadau ym maes deallusrwydd artiffisial yn arafu, felly mae'r DU yn ymrwymedig i sefydlu cydweithrediad rhyngwladol parhaol ar ddiogelwch deallusrwydd artiffisial, gan adeiladu ar sylfeini cytundebau yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial. Er mwyn cadw'r momentwm hwn a sicrhau bod camau yn cael eu cymryd i sicrhau diogelwch deallusrwydd artiffisial, mae Gweriniaeth Korea wedi cytuno i gynnal yr Uwchgynhadledd nesaf ar Ddiogelwch Deallusrwydd Artiffisial ar y cyd â'r DU. Mae Ffrainc wedi cytuno i gynnal yr uwchgynhadledd ddilynol.

---

<sup>84</sup> [UK International Technology Strategy](#), y Swyddfa Dramor, y Gymanwlad a Datblygu, 2023.

<sup>85</sup> [Prime Minister's speech on AI: 26 October 2023](#), Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.

<sup>86</sup> [The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg; Y Swyddfa Dramor, y Gymanwlad a Datblygu; Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.

<sup>87</sup> [World leaders, top AI companies set out plan for safety testing of frontier as first global AI Safety Summit concludes](#), Swyddfa'r Prif Weinidog, 10 Downing Street; Yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

87. Mae Sefydliad Diogelwch Deallusrwydd Artiffisial y DU yn un o'n cyfraniadau allweddol at gydweithrediad rhyngwladol ar ddeallusrwydd artiffisial. Bydd y Sefydliad yn gweithio mewn partneriaeth â gwledydd eraill i hwyluso cydweithrediad rhwng llywodraethau ar drefniadau profi diogelwch deallusrwydd artiffisial a llywodraethu, a datblygu ei alluoedd ei hun. Bydd y Sefydliad yn hwyluso cydweithrediad rhyngwladol mewn tair ffordd allweddol:

- **Partneriaethau:** mae'r Sefydliad Diogelwch Deallusrwydd Artiffisial wedi cytuno i ffurfio partneriaeth â Sefydliad Diogelwch Deallusrwydd Artiffisial yr UD ac â llywodraeth Singapore er mwyn cydweithio ar brofion diogelwch deallusrwydd artiffisial ac mae'n cynnal deialog reolaidd â phartneriaid rhyngwladol ar faterion diogelwch deallusrwydd artiffisial.
- **Adroddiad Rhyngwladol ar Wyddoniaeth Diogelwch Deallusrwydd Artiffisial:**<sup>88</sup> Datgelwyd yr adroddiad yn gyntaf fel yr Adroddiad ar Sefyllfa Gwyddoniaeth yn yr Uwchgyhadledd ar Ddiogelwch Deallusrwydd Artiffisial yn y DU ym mis Tachwedd, lle cytunodd y gwledydd a gynrychiolwyd i ddatblygu adroddiad wedi'i awduru'n rhyngwladol ar alluoedd a risgiau deallusrwydd artiffisial uwch. Yn hytrach na pharatoi deunydd newydd, bydd yn crynhoi'r ymchwil orau sy'n bodoli eisoes ac yn nodi meysydd ymchwil, gan roi synthesis o'r wybodaeth bresennol am risgiau deallusrwydd artiffisial uwch.
- **Cyfnwid Gwybodaeth:** gwerthusiadau ac ymchwil y Sefydliad Diogelwch Deallusrwydd Artiffisial yw'r camau cyntaf i fynd i'r afael â'r bylchau mewn dealltwriaeth rhwng diwydiant, llywodraethau, academyddion, a'r cyhoedd. Bydd hyn yn sicrhau bod y partïon perthnasol, gan gynnwys partneriaid rhyngwladol, yn cael y wybodaeth sydd ei hangen arnynt i lywio'r gwaith o ddatblygu protocolau a rennir.

88. Mae'r DU hefyd yn chwarae rôl ragweithiol drwy nifer o fentrau amlochrog er mwyn bod yn sbardun i'n huchelgais i hyrwyddo'r broses o ddylunio, datblygu, cyflwyno a defnyddio deallusrwydd artiffisial mewn ffordd ddiogel a chyfrifol. Mae hyn yn cynnwys:

- **G7:** Gan weithio mewn cydweithrediad â'n partneriaid yn y fforwm hwn, mae'r DU wedi gwneud cynnydd sylweddol i ymateb yn gyflym i ddatblygiadau technolegol newydd ac ysgogi gwaith ar drefniadau llywodraethu deallusrwydd artiffisial rhyngwladol effeithiol. Ym mis Rhagfyr 2023, o dan Lywyddiaeth Japan, croesawodd Arweinwyr y G7 Fframwaith Polisi Cynhwysfawr Proses Deallusrwydd Artiffisial Hiroshima sy'n cynnwys egwyddorion arweiniol rhyngwladol ar gyfer pawb sy'n ymwneud â deallusrwydd artiffisial a Chod Ymddygiad i sefydliadau sy'n datblygu systemau deallusrwydd artiffisial uwch, yn ogystal â chynllun gwaith i hyrwyddo'r canlyniadau hyn ymhellach.<sup>89</sup> Rydym yn annog y rhai sy'n ymwneud â deallusrwydd artiffisial, ac yn enwedig datblygwyr, i ymgysylltu ymhellach a chefnogi'r canlyniadau hyn. Edrychwn ymlaen at gydweithredu ymhellach ar ddeallusrwydd artiffisial o dan Lywyddiaeth yr Eidal ar Fforwm y G7 yn 2024.
- **G20:** Ym mis Medi 2023, fel rhan o Lywyddiaeth India ar Fforwm y G20, gwnaeth Prif Weinidog y DU gytuno i Ddatganiad Arweinwyr Delhi Newydd a'i gymeradwyo ochr yn ochr ag Aelodau eraill y G20.<sup>90</sup> Gwnaeth y Datganiad ailgadarnhau ymrwymiad y DU i Egwyddorion Deallusrwydd Artiffisial G20 2019 a phwysleisio pwysigrwydd dull llywodraethu sy'n cydbwyso manteision a risgiau deallusrwydd artiffisial ac yn hyrwyddo deallusrwydd artiffisial cyfrifol er mwyn cyflawni Nodau

<sup>88</sup> [International expertise to drive International AI Safety Report](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2024.

<sup>89</sup> [G7 Leaders' Statement on the Hiroshima AI Process](#), Gweinyddiaeth Materion Tramor Japan, 2023.

<sup>90</sup> [G20 New Delhi Leaders' Declaration](#), Gweinyddiaeth Materion Allanol Llywodraeth India, 2023.

Datblygu Cynaliadwy y CU.<sup>91</sup> Bydd y DU yn gweithio'n agos gyda Brasil ar ei huchelgeisiau o ran deallusrwydd artiffisial fel rhan o'i Llywyddiaeth ar G20 2024, a fydd yn canolbwyntio ar ddeallusrwydd artiffisial at ddibenion datblygu cynaliadwy cynhwysol.

- **Y Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial:** Mae'r DU yn parhau i fynd ati i ddylanwadu ar weithgareddau prosiectau rhanddeiliaid y bartneriaeth hon er mwyn llywio'r broses o ddatblygu a defnyddio deallusrwydd artiffisial yn gyfrifol wedi'i seilio mewn hawliau dynol, cynhwysiant, amrywiaeth, arloesedd, a thwf economaidd. Roedd yn bleser gan y DU fod yn bresennol yn Uwchgynhadledd 2023 y bartneriaeth yn Delhi Newydd ym mis Rhagfyr, lle roedd yn cael ei chynrychioli gan y Gweinidog dros Ddeallusrwydd Artiffisial, yr Is-iarll Camrose, a chymeradwyo Datganiad Gweinidogol Delhi Newydd y Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial<sup>92</sup> a chynnal digwyddiad ymylol ar y canlyniadau a'r camau nesaf yn dilyn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial. Mae'r DU hefyd wedi dechrau mandad dwy flynedd fel aelod o'r Pwyllgor Llywio a bydd yn gweithio gyda Chadeiryddiaeth India i sicrhau bod y Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial yn cyrraedd ei llawn botensial.
- **Cyngor Ewrop:** Mae'r DU yn parhau i weithio'n agos gyda gwledydd o'r un anian ar Gonfensiwn arfaethedig Cyngor Ewrop ar ddeallusrwydd artiffisial er mwyn helpu i ddiogelu hawliau dynol, democratiaeth, a rheolaeth y gyfraith. Mae'r Confensiwn yn cynnig cyfle i sicrhau bod y gwerthoedd pwysig hyn yn cael eu codeiddio'n rhyngwladol fel un rhan o ddull gweithredu ehangach o ymdrin â llywodraethu rhyngwladol effeithiol.
- **Y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd:** Mae'r DU yn aelod gweithredol o'r Gweithgor ar Lywodraethu Deallusrwydd Artiffisial ac mae'n cydnabod rôl y fforwm i gefnogi'r broses o roi Egwyddorion Deallusrwydd Artiffisial y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd ar waith<sup>93</sup> a galluogi aelod-wledydd i gyfnewid profiadau ac arferion gorau. Yn 2024, bydd y DU yn cefnogi'r broses o ddiwygio Egwyddorion Deallusrwydd Artiffisial y sefydliad ac yn parhau i gyflwyno astudiaethau achos o Bortffolio'r DU o Dechnegau Sicrwydd Deallusrwydd Artiffisial<sup>94</sup> i'w Gatalog o Adnoddau a Metrigau Adnoddau ar gyfer Deallusrwydd Artiffisial Dibynadwy.<sup>95</sup>
- **Y Cenhedloedd Unedig a'i asiantaethau cysylltiedig:** O ystyried rôl unigryw'r sefydliad i gynnull amrywiaeth eang o wledydd, mae'r DU yn cydnabod gwerth trafodaethau ar ddeallusrwydd artiffisial a arweinir gan y CU ac mae'n ymgysylltu'n rheolaidd i ddylanwadu ar normau byd-eang ynglŷn â deallusrwydd artiffisial. Ym mis Gorffennaf 2023, gwnaeth y DU drefnu a chadeirio sesiwn frifffio gyntaf Cyngor Diogelwch y CU ar ddeallusrwydd artiffisial, a chadeiriodd y Dirprwy Brif Weinidog sesiwn ar risgiau deallusrwydd artiffisial ffin eithaf yn Wythnos Lefel Uchel y CU ym mis Medi 2023. Mae'r DU yn parhau i gydweithio ag amrywiaeth o bartneriaid ar fentrau deallusrwydd artiffisial y CU, gan gynnwys cyd-drafodaethau'r Compact Digidol Byd-eang, sydd â'r nod o hwyluso'r Nodau Datblygu Cynaliadwy drwy dechnolegau megis deallusrwydd artiffisial, monitro'r broses o roi Argymhelliad

<sup>91</sup> [The 17 goals](#), y Cenhedloedd Unedig, 2023.

<sup>92</sup> [GPAI New Delhi Ministerial Declaration](#), y Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial, 2023.

<sup>93</sup> [OECD AI Principles overview](#), y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, 2024.

<sup>94</sup> [CDEI portfolio of AI assurance techniques](#), y Ganolfan Moeseg Data ac Arloesi; Yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>95</sup> [Catalogue of tools and metrics for trustworthy AI](#), y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, dim dyddiad.

UNESCO ar Foeseg Deallusrwydd Artiffisial ar waith,<sup>96</sup> ac ymgysylltu'n adeiladol yn yr Undeb Telathrebu Rhyngwladol, a gynhaliodd yr uwchgynhadledd 'AI for Good' ym mis Gorffennaf 2023. Bydd y DU hefyd yn parhau i weithio'n agos gyda Chorff Cynghori ar Ddeallusrwydd Artiffisial y CU ac mae'n bwrw golwg manwl ar ei adroddiad interim: *Governing AI for Humanity*.<sup>97</sup>

- **Sefydliadau Datblygu Safonau Byd-ean** Mae'r DU yn ymgysylltu'n uniongyrchol â sefydliadau datblygu safonau, megis yr ISO a'r IEC, ac mae'n cefnogi datblygiadau ym maes safonau deallusrwydd artiffisial technegol. Mae'r DU yn hyrwyddo eco-system safonau digidol byd-eang sy'n agored, yn dryloyw ac yn seiliedig ar gonsensws. Mae'r DU hefyd yn anelu at gefnogi arloesedd a chryfhau model a arweinir gan ddiwydiant ac sy'n cynnwys nifer o randdeiliaid er mwyn datblygu safonau deallusrwydd artiffisial technegol, gan gynnwys drwy fentrau megis Hwb Safonau Deallusrwydd Artiffisial y DU.<sup>98</sup> Rydym yn cefnogi rhanddeiliaid yn y DU ac yn eu hannog i gymryd rhan yng ngwaith sefydliadau datblygu safonau er mwyn manteisio ar safonau technegol digidol yma yn y DU a sicrhau safonau technegol digidol sydd wedi cael eu llunio gan werthoedd democrataidd.

89. Hefyd, mae'r DU yn ymrwymedig i sicrhau bod manteision deallusrwydd artiffisial ar gael yn gyffredinol. Mae hyn yn cynnwys gweithio gyda phartneriaid rhyngwladol i ariannu prosiectau deallusrwydd artiffisial diogel a chyfrifol i'w datblygu ledled y byd. Fel y'i cyhoeddwyd yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, mae'r DU yn cyfrannu £38 miliwn drwy ei rhaglen Deallusrwydd Artiffisial ar gyfer Datblygu newydd i gefnogi arloesedd diogel, cyfrifol a chynhwysol ym maes deallusrwydd artiffisial er mwyn cyflymu cynnydd o ran heriau datblygu, gyda'r ffocws cychwynnol yn Affrica.<sup>99</sup> Mae'r swm hwn yn rhan o £80 miliwn o gyllid i hybu rhaglennu deallusrwydd artiffisial er mwyn gwrthsefyll anghydraddoldeb a hyrwyddo ffyniant yn Affrica, lle mae'r DU yn gweithio ochr yn ochr â Chanada, Sefydliad Bill a Melinda Gates, UDA, Google, Microsoft, a phartneriaid Affricanaidd, gan gynnwys Kenya, Nigeria a Rwanda ymhlith eraill.

90. Mae deallusrwydd artiffisial bellach yn elfen hanfodol o'n cydberthnasau dwyochrog ac, mewn rhai achosion, mae'n addas creu partneriaethau dwyochrog dyfnach a mwy ymrwymedig ochr yn ochr ag ymgysylltu amlochrog er mwyn hyrwyddo ein buddiannau cyffredin. Felly, rydym wedi mynd ati i lunio cytundebau dwyochrog mewn meysydd sy'n cynnwys datblygu a chyflwyno deallusrwydd artiffisial mewn ffordd gyfrifol â phartneriaid rhyngwladol allweddol, er mwyn adeiladu sylfaen dros ragor o gydweithio ar lywodraethu deallusrwydd artiffisial. Er enghraifft, fel rhan o Gronfa Partneriaethau Gwyddoniaeth Rhyngwladol yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg,<sup>100</sup> bydd UKRI yn buddsoddi £9 miliwn i ddwyn ynghyd ymchwilwyr ac arloeswyr mewn partneriaethau ymchwil dwyochrog â'r UD. Bydd y partneriaethau hyn yn canolbwyntio ar ddatblygu deallusrwydd artiffisial mwy diogel, cyfrifol a dibynadwy, yn ogystal â deallusrwydd artiffisial at ddibenion gwyddonol. Ers cyhoeddi'r papur gwyn ar reoleiddio deallusrwydd artiffisial ym mis Mawrth 2023 rydym wedi llofnodi:

- **Datganiad yr Iwerydd â'r UD:**<sup>101</sup> sy'n datblygu ein partneriaeth gadarn ynglŷn â deallusrwydd artiffisial, yn seiliedig ar ein gwerthoedd democrataidd cyffredin a'u huchelgais i hyrwyddo arloesi'n ddiogel ac yn gyfrifol ym maes deallusrwydd

<sup>96</sup> [Recommendation on the Ethics of Artificial Intelligence](#), UNESCO, 2023.

<sup>97</sup> [Governing AI for Humanity](#), y Cenedloedd Unedig, 2023.

<sup>98</sup> [The AI Standards Hub](#), Hwb Safonau Deallusrwydd Artiffisial, 2022.

<sup>99</sup> [UK unites with global partners to accelerate development using AI](#), y Swyddfa Dramor, y Gymanwlad a Datblygu, 2023.

<sup>100</sup> [International Science Partnerships Fund \(ISPF\)](#), UKRI, 2023.

<sup>101</sup> [The Atlantic Declaration](#), Swyddfa Prif Weinidog y DU, 10 Downing Street, Swyddfa Dramor, y Gymanwlad a Datblygu, yr Adran Busnes a Masnach, 2023.

artiffisial dros y byd i gyd. Bydd y gwaith sy'n mynd rhagddo o dan Ddatganiad yr Iwerydd, 2023 yn sicrhau y caiff ein cynghrair unigryw ei hatgyfnerthu ar gyfer heriau datblygiadau technolegol newydd.

- **Cytundeb Hiroshima â Japan:**<sup>102</sup> sy'n cynnwys ymrwymiad i ganolbwyntio ar hyrwyddo deallusrwydd artiffisial dynol-ganolog a dibynadwy a'r gallu i ryngweithredu rhwng ein fframweithiau llywodraethu deallusrwydd artiffisial.
- **Cytundeb Downing Street â Gweriniaeth Korea:**<sup>103</sup> sy'n adeiladu ar y cynnydd a wnaed ar ddatblygu deallusrwydd artiffisial diogel a chyfrifol, gan gynnwys yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial – y bydd yr un nesaf yn y gyfres yn cael ei chynnal ar y cyd gan Weriniaeth Korea a'r DU.
- **Datganiad ar y Cyd am Bartneriaeth Strategol â Singapore:**<sup>104</sup> sy'n defnyddio arbenigedd mewn technolegau newydd megis deallusrwydd artiffisial yn y DU a Singapore. Mae'r Adran dros Wyddoniaeth, Arloesedd a Thechnoleg hefyd wedi llofnodi Memorandwm Cyd-ddealltwriaeth ar Dechnolegau Datblygol ym mis Mehefin 2023 ag Awdurdod Datblygu Cyfryngau Cyfathrebu Gwybodaeth Singapore. Yn y Memorandwm Cyd-ddealltwriaeth hwn, cytunodd y ddau barti i gydweithredu ar drefniadau llywodraethu deallusrwydd artiffisial a hwyluso'r broses o ddatblygu systemau sicrwydd deallusrwydd artiffisial effeithiol sy'n gallu rhyngweithredu.

91. Mae gennym nifer o gydberthnasau dwyochrog pwysig eraill ynglŷn â deallusrwydd artiffisial â gwledydd ledled y byd a, lle y bo'n addas, bwriadwn lunio rhagor o gytundebau o'r fath er mwyn cryfhau'r partneriaethau hyn, megis drwy Femoranda Cyd-ddealltwriaeth a Chytundebau Masnach Rydd dwyochrog.

92. Dim ond drwy gydweithio byd-eang effeithiol bydd y DU a'n partneriaid dros y byd yn datgloi'r cyfleoedd a lliniaru'r risgiau sy'n gysylltiedig â deallusrwydd artiffisial. Byddwn yn parhau i ymgysylltu â'n partneriaid rhyngwladol i gefnogi arloesi'n gyfrifol ym maes deallusrwydd artiffisial sy'n ymdrin â'i niweidiau posibl mewn ffordd effeithiol a chymesur ac sy'n gyson â'r egwyddorion a bennwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial. Byddwn hefyd yn gweithio gyda'n gilydd i hyrwyddo cydlyniaeth rhwng ein fframweithiau llywodraethu deallusrwydd artiffisial er mwyn sicrhau y gall busnesau weithredu'n effeithiol yn y DU ac yn y marchnadoedd byd-eang ehangach a bod datblygiadau ym maes deallusrwydd artiffisial yn fuddiol i bobl ledled y byd.

## 5.4. Map ffordd o'n camau nesaf o ran rheoleiddio deallusrwydd artiffisial

93. Yn 2024, byddwn yn:

- Parhau i ddatblygu ein safbwynt polisi domestig ar ddeallusrwydd artiffisial drwy:

---

<sup>102</sup> [The Hiroshima Accord: An enhanced UK-Japan global strategic partnership](#), Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.

<sup>103</sup> [The Downing Street Accord: A United Kingdom-Republic of Korea Global Strategic Partnership](#), Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.

<sup>104</sup> [Joint Declaration by the Prime Ministers of the Republic of Singapore and the United Kingdom of Great Britain and Northern Ireland on a Strategic Partnership](#), Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.



- Ymgysylltu ag amrywiaeth o arbenigwyr ar ymyriadau ar gyfer systemau deallusrwydd artiffisial tra galluog, gan gynnwys cwestiynau ynglŷn â threfniant rhyddhau agored, yn yr haf.
- Cyhoeddi diweddariad ar ein gwaith ar gyfrifoldebau newydd i ddatblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog erbyn diwedd y flwyddyn.
- Cydweithio ym mhob rhan o'r llywodraeth ac â rheoleiddwyr er mwyn dadansoddi ac adolygu bylchau posibl yn y pwerau a'r cylchoedd gwaith rheoleiddio presennol yn barhaus.
- Gweithio'n agos gyda'r Sefydliad Diogelwch Deallusrwydd Artiffisial a fydd yn rhoi dealltwriaeth sylfaenol ar gyfer ein gweithgareddau asesu risg deallusrwydd artiffisial canolog a llywio'r ffordd rydym yn rheoleiddio deallusrwydd artiffisial, yn barhaus. Bydd y Sefydliad Diogelwch Deallusrwydd Artiffisial yn ymateb i reoleiddio risgiau deallusrwydd artiffisial mewn ffordd gymesur sy'n seiliedig ar dystiolaeth.
- Cymryd camau i hyrwyddo cyfleoedd deallusrwydd artiffisial a mynd i'r afael a'i risgiau drwy'r canlynol:
  - Ymgysylltu mewn ffordd wedi'i thargedu ar ein cofrestr risgiau deallusrwydd artiffisial draws-economi a chynllunio i asesu'r fframwaith rheoleiddio o'r gwanwyn ymlaen.
  - Cyhoeddi cais am safbwyntiau yn y gwanwyn er mwyn cael mwy o gyfraniadau ar ein camau nesaf i sicrhau modelau deallusrwydd artiffisial, gan gynnwys Cod Ymarfer posibl ar seiberddiogelwch deallusrwydd artiffisial, yn seiliedig ar ganllawiau'r Ganolfan Seiberddiogelwch Genedlaethol.
  - Cynnal deialog ryngwladol newydd i amddiffyn democratiaeth ac ymdrin â risgiau cyffredin mewn perthynas ag ymyrryd mewn etholiadau cyn yr Uwchgynhadledd nesaf ar Ddiogelwch Deallusrwydd Artiffisial.
  - Lansio cais am dystiolaeth ynglŷn â'r risgiau sy'n gysylltiedig â deallusrwydd artiffisial o ran ymddiriedaeth mewn gwybodaeth a materion cysylltiedig megis ffugiadau dwfn.
  - Ystyried dulliau o sicrhau mwy o dryloywder, gan gynnwys mesurau sy'n ei gwneud yn bosibl i ddeiliaid hawliau ddeall yn well a yw'r cynnwys a gynhyrchir ganddynt yn cael ei ddefnyddio fel mewnbwn mewn modelau deallusrwydd artiffisial.
  - Cyflwyno fesul cam y gofyniad gorfodol ar adrannau llywodraeth ganolog i ddefnyddio'r Safon Cofnodi Tryloywder Algorithmig yn ystod y flwyddyn.
- Ymestyn y swyddogaeth ganolog a chefnogi rheoleiddwyr drwy:
  - Lansio rhaglen newydd gwerth £10 miliwn i helpu rheoleiddwyr i nodi a deall risgiau yn eu priod feysydd a datblygu eu sgiliau a'u dulliau o ymdrin â deallusrwydd artiffisial.
  - sefydlu pwyllgor llywio i gefnogi ac arwain gweithgareddau strwythur cydgysylltu ffurfiol i reoleiddwyr mewn llywodraeth yn ystod y gwanwyn.
  - Gofyn i reoleiddwyr allweddol gyhoeddi diweddariadau ar eu dull gweithredu strategol o ymdrin â deallusrwydd artiffisial erbyn 30 Ebrill.

- Cydweithio â rheoleiddwyr i iteru ac ehangu ar ein canllawiau traws-sector cychwynnol ar roi'r egwyddorion ar waith, gyda diweddariadau pellach yn yr arfaeth erbyn yr haf.
- Annog diwydiant, arloeswyr a chyflogeion i fabwysiadu deallusrwydd artiffisial mewn ffordd effeithiol a rhoi cymorth iddynt drwy:
  - Lansio Hwb Deallusrwydd Artiffisial a Digidol peilot ar y cyd â'r Fforwm Cydweithredu Rheoleiddio Digidol yn y gwanwyn.
  - Cyhoeddi *Introduction to AI Assurance* yn y gwanwyn.
  - Cyhoeddi canllawiau wedi'u diweddarau ar ddefnyddio deallusrwydd artiffisial ym maes adnoddau dynol a recriwtio yn y gwanwyn.
  - Lansio cynllun Hanfodion Rheoli Deallusrwydd Artiffisial i bennu safon arfer da gofynnol i gwmnïau sy'n gwerthu cynhyrchion a gwasanaethau deallusrwydd artiffisial erbyn diwedd y flwyddyn.
  - Cyhoeddi diweddariad ar ein canllaw ar brosesau sy'n cael ei ddatblygu erbyn diwedd y flwyddyn.
- Cefnogi cydweithrediad rhyngwladol ynglŷn â threfniadau llywodraethu deallusrwydd artiffisial drwy'r canlynol:
  - Rhoi ein partneriaeth gwerth £9 miliwn newydd ei chyhoeddi â'r UD ynglŷn â deallusrwydd artiffisial cyfrifol ar waith fel rhan o Gronfa Partneriaethau Gwyddoniaeth Rhyngwladol yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg.
  - Cyhoeddi'r iteriad cyntaf o'r Adroddiad Rhyngwladol ar Wyddoniaeth Diogelwch Deallusrwydd Artiffisial yn y gwanwyn.
  - Rhannu gwybodaeth newydd â phartneriaid rhyngwladol drwy'r Sefydliad Diogelwch Deallusrwydd Artiffisial yn barhaus.
  - Rhoi cymorth i Weriniaeth Korea a Ffrainc ar yr Uwchgynadleddau nesaf ar Ddiogelwch Deallusrwydd Artiffisial yn barhaus, ac ystyried rôl bosibl Uwchgynadleddau ar Ddiogelwch Deallusrwydd Artiffisial y tu hwnt i'r rhain.
  - Parhau â phartneriaethau dwyochrog ac amlochrog ynglŷn â deallusrwydd artiffisial, gan gynnwys fforwm y G7, fforwm y G20, Cyngor Ewrop, y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, y Cenhedloedd Unedig, a'r Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial.

## 6. Crynodeb o dystiolaeth yr ymgynghoriad ac ymateb y llywodraeth

94. Mae'r bennod hon yn rhoi crynodeb o'r dystiolaeth ysgrifenedig a gyflwynwyd i ni mewn ymateb i'n hymgynghoriad ac wedyn ymateb y llywodraeth. Mae'r bennod hon wedi'i strwythuro ar sail y 10 categori a ddefnyddiwyd gennym i drefnu ein 33 o gwestiynau yn yr ymgynghoriad:

- Yr egwyddorion deallusrwydd artiffisial traws-sector diwygiedig.
- Dyletswydd statudol i roi sylw.
- Swyddogaethau canolog newydd i gefnogi'r fframwaith.
- Monitro a gwerthuso'r fframwaith.
- Galluoedd rheoleiddwyr.
- Adnoddau ar gyfer deallusrwydd artiffisial dibynadwy.
- Syniadau i gloi.
- Cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial.
- Modelau sylfaen a'r fframwaith rheoleiddio.
- Pyllau tywod a meiciau arbrofi deallusrwydd artiffisial.

95. Cawsom gyfanswm o 409 o ymatebion ysgrifenedig i'r ymgynghoriad gan sefydliadau ac unigolion. Mae Atodiad A yn rhoi trosolwg o bwy a gyflwynodd yr ymatebion ac yn amlinellu ein dull o'u dadansoddi. Gwnaethom hefyd fynd ati i ymgysylltu â 364 o unigolion drwy ddigwyddiadau bord gron, gweithdai technegol, cyfarfodydd dwyochrog, a rhaglen o ymgysylltu parhaus â rheoleiddwyr. Er ein bod yn cydblethu dirnadaethau o'r gweithgarwch ymgysylltu hwn drwy'r dadansoddiad cyfan, mae Atodiad A yn rhoi trosolwg manwl o ganfyddiadau ein gweithgarwch ymgysylltu.

### 6.1. Yr egwyddorion deallusrwydd artiffisial traws-sector diwygiedig

- 1. A ydych chi'n cytuno y byddai gofyniad ar sefydliadau i'w gwneud yn glir eu bod yn defnyddio deallusrwydd artiffisial yn sicrhau gwell tryloywder?**
- 2. A oes unrhyw fesurau eraill y gallem eu gosod ar sefydliadau i wella tryloywder o ran deallusrwydd artiffisial?**
- 3. A ydych chi'n cytuno bod y llwybrau presennol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch yn ddigonol?**
- 4. A allai'r llwybrau presennol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch gael eu gwella, os o gwbl.**

**3. A ydych chi'n cytuno y bydd yr egwyddorion traws-sector diwygiedig, unwaith y cânt eu rhoi ar waith yn effeithiol, yn ymdrin â'r risgiau sy'n codi o ganlyniad i dechnolegau deallusrwydd artiffisial?**

**4. Beth, os unrhyw beth, sydd ar goll yn yr egwyddorion diwygiedig?**

Crynodeb o gwestiynau 1-6:

96. Roedd dros hanner yr ymatebwyr yn cytuno y byddai'r egwyddorion diwygiedig, o'u rhoi ar waith yn effeithiol, yn ymdrin â'r risgiau allweddol sy'n codi o ganlyniad i dechnolegau deallusrwydd artiffisial. Roedd yr egwyddorion diwygiedig yn cynnwys diogelwch a chadernid; tryloywder ac esboniadwyedd priodol; tegwch; atebolrwydd a llywodraethu; y gallu i herio a cheisio iawn. Fodd bynnag, roedd ymatebwyr hefyd yn dadlau o blaid cynnwys hawliau dynol yn benodol, cadernid gweithredol, ansawdd data, cysondeb rhyngwladol, risgiau systemig ac effeithiau cymdeithasol ehangach, cynaliadwyedd, ac addysg a llythrennedd.

97. Roedd ymatebwyr am weld rhagor o fanylion am sut y câi'r egwyddorion eu rhoi ar waith, galluoedd rheoleiddwyr, ac unrhyw ryngweithio â chyfreithiau sy'n bodoli eisoes. Pwysleisiodd ymatebwyr yn gyson pa mor gyflym y mae technoleg yn newid, gan nodi y dylai'r fframwaith fod yn un y gellir ei addasu a dylai gael ei gefnogi gan weithgarwch monitro a gwerthuso. Roedd rhai ymatebwyr yn pryderu na fyddai modd gorfodi'r egwyddorion yn ddigonol, gan nodi diffyg statws statudol.

98. Cafwyd cefnogaeth gref i nifer o fesurau tryloywder ymhlith ymatebwyr. Pwysleisiodd ymatebwyr fod tryloywder yn allweddol o ran meithrin ymddiriedaeth y cyhoedd, atebolrwydd a fframwaith rheoleiddio effeithiol a dilysadwy. Roedd y mwyafrif o'r ymatebwyr yn cytuno y byddai gofyniad ar sefydliad i'w gwneud yn glir pan oeddent yn defnyddio deallusrwydd artiffisial yn gwella tryloywder. Roedd y rhai a anghytunodd yn teimlo y byddai labelu'r defnydd o ddeallusrwydd artiffisial naill ai'n annigonol neu'n anghymesur o feichus. Awgrymodd ymatebwyr amrywiaeth o fesurau tryloywder, gan gynnwys datgelu mewnbynnau fel cyfrifiannu a data yn gyhoeddus; labelu'r defnydd o ddeallusrwydd artiffisial ac allbynnau; dewisiadau optio i mewn a dewisiadau dynol yn lle prosesu awtomataidd; esboniadau ynglŷn â chanlyniadau, effeithiau a chyfyngiadau deallusrwydd artiffisial; cofrestrau deallusrwydd artiffisial cyhoeddus neu sefydliadol; datgelu manylion modelau i reoleiddwyr; ac adnoddau sicrwydd annibynnol, gan gynnwys archwiliadau a safonau technegol.

99. Nododd y rhan fwyaf o'r ymatebwyr nad yw'r llwybrau presennol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch drwy'r fframweithiau cyfreithiol presennol yn ddigonol. Nododd ymatebwyr y gall fod yn anodd nodi niweidiau sy'n gysylltiedig â deallusrwydd artiffisial a bod costau uchel ymglyfreitha yn aml yn atal unigolion rhag ceisio iawn. Roedd llawer o'r ymatebwyr am weld y llywodraeth yn egluro'r hawliau a'r cyfrifoldebau cyfrifol sy'n gysylltiedig â deallusrwydd artiffisial, gyda nifer yn awgrymu y dylid gwneud hyn drwy ganllawiau rheoleiddiol. Roedd rhai o blaid cyflwyno gofynion statudol. Argymhellodd ymatebwyr y dylid sefydlu llwybrau hygyrch i geisio iawn, gyda rhai yn dadlau o blaid system iawn draws-sector, ganolog megis ombwdsmon penodedig ar gyfer deallusrwydd artiffisial. Nododd ymatebwyr hefyd y byddai angen cytundebau rhyngwladol er mwyn sicrhau llwybrau effeithiol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial dros ffiniau neu geisio iawn yn eu cylch. Pwysleisiodd ymatebwyr y byddai mwy o dryloywder ynglŷn â deallusrwydd artiffisial yn helpu i'w gwneud yn haws ceisio iawn am amrywiaeth eang o niweidiau posibl, gan gynnwys ymyrryd â hawliau eiddo deallusol.

## Ymateb:

100. Mae'r llywodraeth am sicrhau bod y DU yn cadw ei statws fel arweinydd byd-eang ym maes deallusrwydd artiffisial. Mae hyn yn golygu hyrwyddo ffyrdd diogel a chyfrifol o arloesi er mwyn sicrhau ein bod yn sicrhau'r manteision mwyaf posibl y gall deallusrwydd artiffisial eu cynnig ledled y wlad. Mae ein hegwyddorion traws-sector yn nodi ein disgwyliadau o ran dylunio, datblygu a chymhwyso deallusrwydd artiffisial yn gyfrifol er mwyn helpu i roi arweiniad i fusnesau a sefydliadau sy'n creu ac yn defnyddio'r technolegau hyn. Mae'n galonogol gweld bod y rhan fwyaf o'r ymatebwyr yn cytuno y bydd yr egwyddorion traws-sector diwygiedig yn ymdrin â'r risgiau sy'n codi o ganlyniad i ddeallusrwydd artiffisial pan gânt eu rhoi ar waith yn effeithiol.

101. Rydym yn disgwyl i reoleiddwyr gymhwyso'r egwyddorion o dan eu cylchoedd gwaith cyfredol ac yn unol â'n cyfreithiau a'n gwerthoedd presennol, gan barchu hanes hir y DU o ddemocratiaeth, rheolaeth gadarn y gyfraith, ac ymrwymadau i hawliau dynol a chynaliadwyedd amgylcheddol. Gan fod agweddau ar y gwerthoedd a'r rheolau hyn wedi'u hymgorffori yn gyfraith y mae rhwymedigaeth ar reoleiddwyr i'w dilyn, yn ein barn ni, nid oes angen cynnwys democratiaeth, hawliau dynol, rheolaeth y gyfraith, na chynaliadwyedd yn benodol o fewn yr egwyddorion eu hunain. Bydd y canllawiau rydym yn eu cyhoeddi ochr yn ochr â'r ymateb hwn i'r ymgynghoriad yn helpu rheoleiddwyr i roi'r egwyddorion ar waith yn eu priod feysydd.

102. Mae'r egwyddorion eisoes yn ymdrin â materion a godwyd gan ymatebwyr mewn cysylltiad â chadernid gweithredol (diogelwch a chadernid) a diogelu data (thryloywder, tegwch, ac atebolrwydd). Rydym yn disgwyl i bawb yn ystod cylch oes deallusrwydd artiffisial gadw at y fframweithiau cyfreithiol presennol, gan gynnwys cyfraith diogelu data. Mae deddfwriaeth diogelu data cyfredol y DU (GDPR y DU a Deddf Diogelu Data 2018) yn rheoleiddio'r ffordd y caiff systemau deallusrwydd artiffisial a thechnolegau eraill eu datblygu pan fo data personol ynghlwm wrth hynny. Bydd y Bil Diogelu Data a Gwybodaeth Ddigidol yn egluro hawliau testunau data i fesurau diogelu penodol pan fyddant yn destun penderfyniadau awtomataidd yn unig a fydd yn effeithio arnynt. At hynny, mae Swyddfa'r Comisiynydd Gwybodaeth wedi creu canllawiau penodol ynglŷn â sut i ddefnyddio data ar gyfer deallusrwydd artiffisial a chydymffurfio â chyfraith diogelu data.<sup>105</sup> Y tu hwnt i gwmpas cyfraith diogelu data, mae'r llywodraeth yn asesu nifer o ymyriadau posibl sy'n gyson â'r egwyddorion fel rhan o'n gwaith o annog ffyrdd cyfrifol a diogel o ddatblygu deallusrwydd artiffisial tra galluog. Er enghraifft, rydym yn ystyried os a sut y dylid cyflwyno mesurau sydd wedi'u targedu at ddatblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog mewn perthynas â gofynion tryloywder (er enghraifft, ar ddata hyfforddi), rheoli risg, a rhwymedigaethau ynglŷn ag atebolrwydd a llywodraethu corfforaethol. Yn yr un modd, bydd ein gweithgareddau rheoli risg canolog yn nodi ac yn monitro nifer o risgiau, gan gynnig trosolwg traws-economi a fydd yn cofnodi risgiau systemig ac effeithiau ehangach ar gymdeithas.

103. Rydym yn cydnabod y gefnogaeth gyffredinol i dryloywder a byddwn yn parhau â'r gwaith i asesu a yw mesurau yn cynnig y tryloywder mwyaf ystyrion i ddefnyddwyr terfynol deallusrwydd artiffisial a'r rhai sy'n rhan o gylch oes deallusrwydd artiffisial, a pha fesurau. Mae'n bwysig ein bod yn gweithredu ar sail tystiolaeth o ran tryloywder. Mae'r Safon Cofnodi Tryloywder Algorithmig yn ffordd ymarferol o sicrhau tryloywder a ddatblygwyd drwy ymgysylltu â'r cyhoedd ac mae wedi cael ei threialu ledled y DU.<sup>106</sup> Mae'r safon yn helpu sefydliadau yn y sector cyhoeddus i roi gwybodaeth glir am yr adnoddau algorithmig a ddefnyddir ganddynt i wneud penderfyniadau. Fel y crybwyllir yn adran 5.1, byddwn yn

<sup>105</sup> [Guidance on AI and data protection](#), Swyddfa'r Comisiynydd Gwybodaeth, 2023.

<sup>106</sup> Datblygwyd gan yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg a'r Swyddfa Digidol a Data Ganolog i'r sector cyhoeddus.

ei gwneud yn ofynnol i bob un o adrannau'r llywodraeth ddefnyddio'r safon a bwriadwn ymestyn hyn i'r sector cyhoeddus yn fwy cyffredinol dros amser. Er y gall mesurau fel dyfrnodau helpu defnyddwyr i nodi cynnwys a gynhyrchir drwy ddeallusrwydd artiffisial, mae angen i ni sicrhau bod ymyriadau a gynigir yn gadarn, na all eu diystyru'n hawdd, a'u bod yn sicrhau canlyniadau cadarnhaol. Er mwyn sicrhau mwy o dryloywder ynglŷn ag allbynnau deallusrwydd artiffisial, cyhoeddwyd dogfen "Emerging processes for frontier AI safety" gennym. Mae'n amlinellu tri maes ymarfer sy'n gysylltiedig â nodi cynnwys a gynhyrchir drwy ddeallusrwydd artiffisial, gan gynnwys technegau ymchwil, dyfrnodau, a chronfeydd data allbwn deallusrwydd artiffisial.<sup>107</sup> Fel y crybwyllir yn adran 5.2.2 byddwn yn diweddarau'r canllaw hwn erbyn diwedd y flwyddyn ac yn parhau i annog cwmnïau deallusrwydd artiffisial i ddatblygu arferion gorau.

104. Mae ein rheoleiddwyr arbenigol eisoes yn defnyddio eu cylchoedd gwaith presennol i roi'r egwyddorion deallusrwydd artiffisial ar waith, gan gynnwys egwyddor herio a cheisio iawn sy'n cynnwys disgwyliadau ynglŷn ag egluro'r llwybrau presennol i geisio iawn. Rydym yn cydnabod y cysylltiad rhwng neilltuo atebolrwydd mewn ffordd deg ac effeithiol drwy gydol oes deallusrwydd artiffisial ac argaeledd ac eglurder y llwybrau i geisio iawn. Mae ein gwaith ar ystyried y fframweithiau atebolrwydd presennol ac atebolrwydd drwy'r gadwyn werth yn mynd rhagddo ac mae'n cynnwys dadansoddi bodolaeth ffyrdd o geisio iawn. Fel cam cyntaf tuag at sicrhau bod atebolrwydd yn cael ei neilltuo mewn ffordd deg ac effeithiol, mae'r llywodraeth yn ystyried cyflwyno gofynion cyfrwymol sydd wedi'u targedu at ddatblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog a all olygu creu neu neilltuo pwerau rheoleiddio newydd.

## 6.2. Dyletswydd statudol i roi sylw

**7. A ydych chi'n cytuno y byddai cyflwyno dyletswydd statudol ar reoleiddwyr i roi sylw priodol i'r egwyddorion yn egluro ac yn cryfhau mandadau rheoleiddwyr i roi ein hegwyddorion ar waith, a chadw ymagwedd hyblyg tuag at eu rhoi ar waith?**

**8. A oes ymyriad statudol amgen a fyddai'n fwy effeithiol?**

Crynodeb o gwestiynau 7-8:

105. Roedd y rhan fwyaf o'r ymatebwyr yn cytuno i ryw raddau neu'n cytuno'n gryf y byddai cyflwyno dyletswydd statudol ar reoleiddwyr i roi sylw priodol i'r egwyddorion a amlinellwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial yn egluro ac yn cryfhau mandadau rheoleiddwyr i roi'r egwyddorion ar waith, a chadw ymagwedd hyblyg tuag at eu rhoi ar waith. Fodd bynnag, nododd bron chwarter y byddai angen adnoddau a galluoedd ychwanegol ar reoleiddwyr er mwyn iddynt ymgymryd â dyletswydd statudol yn effeithiol.

106. Dadleuodd tua thraean o'r ymatebwyr y byddai angen mesurau statudol, ychwanegol wedi'u targedu i roi'r fframwaith rheoleiddio ar waith yn effeithiol. Awgrymodd llawer y dylid ehangu pwerau rheoleiddwyr, gan nodi y byddai cylch gwaith statudol presennol rhai o'r rheoleiddwyr yn cyfyngu ar eu gallu i roi'r fframwaith ar waith. Yn benodol, cododd ymatebwyr yr angen i adolygu pwerau a galluoedd ymchwiliol rheoleiddwyr o ran deallusrwydd artiffisial a'u hehangu o bosibl.

<sup>107</sup> [Emerging Processes for Frontier AI Safety](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

107. Dadleuodd rhai o blaid mesurau statudol llorweddol ehangach, megis deddfwriaeth benodol ynglŷn â deallusrwydd artiffisial, rheoleiddiwr newydd ar gyfer deallusrwydd artiffisial, a rheolau llym ynglŷn â'r defnydd o ddeallusrwydd artiffisial mewn cyd-destunau penodol.

108. Roedd ymatebwyr eraill yn teimlo y gallai dyletswydd i roi sylw, pe bai'n cael ei rhoi ar waith yn fryslog, darfu ar drefniadau rheoleiddio, arloesi, ac ymddiriedaeth. Argymhellodd yr ymatebwyr hyn y dylai'r ddyletswydd gael ei hadolygu ar ôl cyfnod o weithredu anstatudol, yn enwedig i weld a oedd unrhyw ryngweithio â chyfreithiau a chylchoedd gwaith rheoleiddio sy'n bodoli eisoes. Nododd rhai ymatebwyr nad oedd nod terfynol na therfynau amser y fframwaith rheoleiddio deallusrwydd artiffisial yn glir, a bod hynny'n achosi ansicrwydd.

## Ymateb:

109. Mae'n galonogol bod yr ymatebwyr i'r cwestiwn hwn yn frwd o blaid rhoi ein hegwyddorion deallusrwydd artiffisial traws-sector ar waith mewn ffordd briodol ac effeithiol Rydym yn croesawu'r gefnogaeth eang i ddyletswydd statudol ar reoleiddwyr, gan gydnabod bod ymatebwyr hefyd wedi cynnig amodau a dulliau amgen y gellid eu defnyddio i roi'r fframwaith ar waith yn effeithiol. Fel y'i nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, disgwyliwn gyflwyno dyletswydd statudol ar reoleiddwyr yn ei gwneud yn ofynnol iddynt roi sylw priodol i'r egwyddorion ar ôl adolygu cyfnod cychwynnol o'u rhoi ar waith yn anstatudol.

110. Rydym yn cydnabod pryderon ymatebwyr y gallai rhoi dyletswydd i roi sylw ar waith mewn ffordd fryslog darfu ar arloesi'n gyfrifol ym maes deallusrwydd artiffisial. Ni fyddwn yn brysio i ddeddfu ond byddwn yn gwerthuso a yw'n angenrheidiol ac yn effeithiol cyflwyno dyletswydd statudol i roi sylw priodol i'r egwyddorion ar reoleiddwyr. Ar hyn o bryd rydym o'r farn bod dull gweithredu anstatudol yn cynnig hyblygrwydd hollbwysig ond byddwn yn parhau i adolygu hyn, er enghraifft, drwy asesu'r diweddariadau ar ddulliau strategol o ymdrin â deallusrwydd artiffisial y mae'r llywodraeth wedi gofyn i nifer o reoleiddwyr eu cyhoeddi erbyn 30 Ebrill 2024. Byddwn hefyd yn gweithio gydag adrannau'r llywodraeth a rheoleiddwyr i ddadansoddi ac adolygu bylchau posibl yn y pwerau a'r cylchoedd gwaith rheoleiddio presennol.

111. Mae'n bleser gennym weld bod llawer o'r rheoleiddwyr yn cymryd camau rhagweithiol i ymdrin â deallusrwydd artiffisial a rhoi'r egwyddorion ar waith o fewn eu cylchoedd gwaith. Mae hyn yn cynnwys gwaith gan yr Awdurdod Cystadleuaeth a Marchnadoedd, yr Awdurdod Safonau Hysbysebu, a'r Swyddfa Gyfathrebiadau (Ofcom).<sup>108</sup> Mae eraill yn datblygu eu cynlluniau presennol mewn ffyrdd sy'n gyson â'r egwyddorion hyn, megis Swyddfa'r Comisiynydd Gwybodaeth, a'r Asiantaeth Rheoleiddio Meddyginiaethau a Chynhyrchion Gofal Iechyd.<sup>109</sup>

112. Rydym yn parhau i weithio'n agos gyda rheoleiddwyr i ddatblygu'r fframwaith, sicrhau ei fod yn cael ei roi ar waith mewn ffordd gydlynol, a meithrin galluedd rheoleiddwyr. Er mwyn cefnogi dull gweithredu cydlynol rhwng sectorau, rydym yn cyhoeddi canllawiau cychwynnol i reoleiddwyr, ochr yn ochr â'r ymateb hwn ynglŷn â sut i gymhwysu'r egwyddorion deallusrwydd artiffisial traws-sector o fewn eu cylchoedd gwaith presennol. Byddwn yn diweddarau'r canllawiau hyn dros amser er mwyn sicrhau eu bod yn adlewyrchu

---

<sup>108</sup> [AI Foundation Models: initial review](#), yr Awdurdod Cystadleuaeth a Marchnadoedd, 2023; [Generative AI & Advertising: Decoding AI Regulation](#), yr Awdurdod Safonau Hysbysebu, 2023; [What generative AI means for the communications sector](#), Ofcom, 2023.

<sup>109</sup> [How do we ensure fairness in AI?](#), Swyddfa'r Comisiynydd Gwybodaeth, 2023; [Software and Artificial Intelligence \(AI\) as a Medical Device](#), yr Asiantaeth Rheoleiddio Meddyginiaethau a Chynhyrchion Gofal Iechyd, a ddiweddarwyd yn 2023 [2021].

datblygiadau yn ein cyfundrefn a datblygiadau technolegol ym maes deallusrwydd artiffisial. Byddwn yn sefydlu pwyllgor llywio erbyn gwanwyn 2024 i gefnogi a thywys gweithgarwch swyddogaeth gydgyssylltu ganolog i reoleiddwyr (gweler adran 5.1.2 am fanylion).

113. Nodwn bryderon ymatebwyr yn yr ymgynghoriad na ddylai unrhyw reolau newydd ar gyfer deallusrwydd artiffisial wrthddweud na dyblygu cyfreithiau presennol. Byddwn yn parhau i werthuso unrhyw fylchau neu wrthdaro posibl o fewn cylchoedd gwaith statudol presennol rheoleiddwyr a'r fframweithiau deddfwriaethol presennol. Yn y papur gwyn, dywedwyd y byddem yn parhau i adolygu tirwedd ehangach deallusrwydd artiffisial er mwyn llywio iteriadau ar y fframwaith rheoleiddio yn y dyfodol, gan gynnwys a fydd angen ymyriadau pellach o ran modelau sylfaen. Byddwn yn ymgynghori ar ein cynllun ar gyfer monitro a gwerthuso'r fframwaith rheoleiddio yn 2024 (gweler ein hymateb i gwestiynau ynglŷn â monitro a gwerthuso yn adran 6.4 am ragor o fanylion).

### 6.3. Swyddogaethau canolog newydd i gefnogi'r fframwaith

**9. A ydych chi'n cytuno y byddai'r swyddogaeth a amlinellwyd yn adran 3.3.1 yn fuddiol o ran ein fframwaith rheoleiddio deallusrwydd artiffisial pe bai'n cael ei chyflawni'n ganolog?**

**10. Beth, os unrhyw beth, sydd ar goll yn swyddogaethau canolog?**

**11. A wyddoch am unrhyw sefydliadau sydd eisoes yn bodoli a ddylai gyflawni un neu fwy o'n swyddogaethau canolog arfaethedig?**

**12. A oes gweithgareddau ychwanegol a fyddai'n helpu busnesau i arloesi a defnyddio technolegau deallusrwydd artiffisial yn hyderus?**

**12.1. Os felly, a ddylai'r gweithgareddau hyn gael eu cyflawni gan y llywodraeth, rheoleiddwyr, neu sefydliad gwahanol?**

**13. A oes gweithgareddau ychwanegol a fyddai'n helpu unigolion a defnyddwyr i ddefnyddio technolegau deallusrwydd artiffisial yn hyderus?**

**13.1 Os felly, a ddylai'r gweithgareddau hyn gael eu cyflawni gan y llywodraeth, rheoleiddwyr, neu sefydliad gwahanol?**

**14. Sut y gallwn osgoi sefyllfa lle mae canllawiau ar ddeallusrwydd artiffisial a gyhoeddir gan wahanol reoleiddwyr yn gorgyffwrdd, yn dyblygu, neu'n gwrthddweud ei gilydd?**

Crynodeb o gwestiynau 9-14:

114. Cytunodd bron bob ymatebydd y byddai'n fuddiol i'r fframwaith rheoleiddio deallusrwydd artiffisial pe bai'r swyddogaethau arfaethedig yn cael eu cyflawni'n ganolog, gyda llawer yn canmol y dull gweithredu am sicrhau bod y llywodraeth yn gallu monitro ac iteru'r fframwaith.

115. Er bod yr ymatebwyr yn cefnogi'r swyddogaethau canolog arfaethedig yn gyffredinol, roedd llawer am gael rhagor o fanylion ar bob swyddogaeth a'i gweithgareddau. Roedd rhai ymatebwyr yn teimlo y dylai fod mwy o bwyslais ar bartneriaethau a chydweithio i gyflawni'r gweithgareddau. Roedd ymatebwyr hefyd am weld mwy o fanylion ar gydweithio



rhyngwladol. Awgrymodd rhai y dylai'r llywodraeth roi blaenoriaeth ar greu'r swyddogaeth risg ganolog. Ymhlith yr ymatebion hyn, nododd ychydig y dylid rhoi mwy o ystyriaeth i risgiau moesegol a chymdeithasol.

116. Pwysleisiodd ymatebwyr y dylai'r swyddogaethau rheoleiddio gael eu hadeiladu ar gryfderau presennol tirwedd reoleiddio'r DU, gyda thua thraean yn nodi rheoleiddwyr fel sefydliadau a ddylai gyflawni un neu fwy o'r swyddogaethau canolog. Ar y cyfan, pwysleisiodd ymatebwyr y byddai angen cydweithio rhwng y llywodraeth, rheoleiddwyr, diwydiant, cymdeithas sifil, y byd academiaidd, a'r cyhoedd er mwyn cyflawni'r swyddogaethau'n effeithiol. Roedd dros chwarter yr ymatebwyr yn teimlo y gallai sefydliadau ymchwil a melinau trafod sy'n canolbwyntio ar dechnoleg helpu i gyflawni'r swyddogaethau canolog.

117. Awgrymodd ymatebwyr nifer o weithgareddau ychwanegol y gallai'r llywodraeth a'r rheoleiddwyr eu cynnig er mwyn cefnogi diwydiant. Roedd tua thraean o'r ymatebwyr yn teimlo y byddai cynhyrchion hyfforddi ac adnoddau addysgol yn helpu sefydliadau i gymhwyso'r egwyddorion at arferion busnes pob dydd. Awgrymodd bron chwarter y dylai rheoleiddwyr baratoi canllaw er mwyn galluogi busnesau i arloesi'n hyderus. Nododd rhai bwysigrwydd fframweithiau y gellir eu gweithredu'n rhyngwladol i reoleiddio deallusrwydd artiffisial er mwyn sicrhau baich cydymffurfio bach ar sefydliadau sy'n adeiladu, yn gwerthu, ac yn defnyddio technolegau deallusrwydd artiffisial. Dadleuodd ymatebwyr hefyd fod angen gwneud rhagor o waith er mwyn sicrhau bod data amrywiol, ansawdd uchel sy'n cael eu cyrchu mewn ffordd foesegol ar gael i fusnesau er mwyn cefnog eu hymdrechion i arloesi ym maes deallusrwydd artiffisial.

118. Wrth feddwl am weithgareddau ychwanegol i unigolion a defnyddwyr, rhoddodd ymatebwyr flaenoriaeth ar sicrhau tryloywder o'r egwyddorion traws-sector, gyda bron hanner yn dadlau y dylai fod modd i unigolion a defnyddwyr nodi pryd a sut mae deallusrwydd artiffisial yn cael ei ddefnyddio gan wasanaeth neu sefydliad. Roedd dros draean o'r ymatebwyr yn teimlo y byddai addysg a hyfforddiant yn galluogi defnyddwyr i ddefnyddio cynhyrchion a gwasanaethau deallusrwydd artiffisial mewn ffordd ddiogel a mwy effeithiol.

119. Awgrymodd tua thraean mai'r swyddogaethau canolog arfaethedig fyddai'r ffordd fwyaf effeithiol o osgoi canllawiau sy'n gorgyffwrdd, yn dyblygu, neu'n gwrthddweud ei gilydd.

## Ymateb:

120. Rydym yn croesawu'r gefnogaeth gref i'r swyddogaethau canolog a gynigiwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial er mwyn cydgysylltu, monitro, ac addasu'r fframwaith deallusrwydd artiffisial. Gyda'i gilydd, bydd y swyddogaethau hyn yn rhoi eglurder, yn sicrhau bod y fframwaith yn gweithio yn ôl y bwriad, ac yn diogelu dull rheoleiddio'r DU rhag datblygiadau yn y dyfodol. Dyna pam rydym eisoes wedi dechrau sefydlu'r swyddogaeth ganolog o fewn y llywodraeth er mwyn ymgymryd â'r gweithgareddau a gynigiwyd yn y papur gwyn (gweler adran 5.1.2 am fanylion).

121. Nodwn bryderon ymatebwyr ynglŷn â'r risgiau posibl sy'n codi o ganlyniad i'r datblygiadau cyflym ym maes technoleg deallusrwydd artiffisial. Rydym eisoes wedi pennu gweithgareddau monitro ac asesu risg y swyddogaeth ganolog yn yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, sy'n adlewyrchu'r argymhelliad cryf gan ymatebwyr i roi prosesau rheoli risgiau deallusrwydd artiffisial traws-economi ar waith fel blaenoriaeth. Bydd ein gweithgareddau asesu risgiau canolog yn nodi, yn mesur, ac yn monitro risgiau deallusrwydd artiffisial sy'n bodoli eisoes ac sy'n dod i'r amlwg drwy ddefnyddio arbenigedd o bob rhan o'r llywodraeth, diwydiant, a'r byd academiaidd, gan gynnwys y Sefydliad

Diogelwch Deallusrwydd Artiffisial. Bydd hyn yn ein galluogi i fonitro risgiau yn gyfannol a nodi unrhyw fylchau posibl yn ein dull gweithredu. Bydd sganio'r gorwel yn estyn ein gweithgareddau asesu risgiau canolog, gan fonitro tueddiadau a chyfleoedd sy'n dod i'r amlwg ym maes deallusrwydd artiffisial er mwyn manteisio i'r eithaf arno a hefyd ymdrin â'i risgiau mewn ffordd gymesur. Eleni, byddwn yn cynnal gweithgarwch ymgysylltu wedi'i dargedu ar ein cofrestr risgiau deallusrwydd artiffisial draws-economi.

122. Gan adlewyrchu safbwyntiau ymatebwyr y bydd y swyddogaeth ganolog arfaethedig yn helpu rheoleiddwyr i osgoi paratoi canllawiau sy'n gorgyffwrdd, sy'n dyblygu, neu'n gwrthddefn ei gilydd, rydym yn datblygu swyddogaeth gydgyssylltu er mwyn helpu rheoleiddwyr i ddehongli a chymhwyso'r egwyddorion o fewn eu cylchoedd gwaith (gweler adran 5.1.2 am fanylion). Fel rhan o'r broses hon, byddwn yn sefydlu pwyllgor llywio yn y gwanwyn a fydd yn cynnwys cynrychiolwyr o'r llywodraeth a rheoleiddwyr allweddol i'w helpu i gyfnewid gwybodaeth a chydgyssylltu ynglŷn â llywodraethu deallusrwydd artiffisial. Er mwyn cefnogi rheoleiddwyr ymhellach a sicrhau y manteisir i'r eithaf ar gryfder y DU ym maes ymchwil i ddeallusrwydd artiffisial yn ein fframwaith rheoleiddio, rydym hefyd wedi cyhoeddi pecyn gwerth £10 miliwn i gefnogi galluoedd deallusrwydd artiffisial rheoleiddwyr ac ymrwymiad newydd gan Ymchwil ac Arloesi yn y DU (UKRI) i wella cysylltiadau rhwng rheoleiddwyr a'r sgiliau, yr arbenigedd, a'r gweithgareddau a gefnogir gan ei fuddsoddiad mewn ymchwil i ddeallusrwydd artiffisial yn y dyfodol.

123. Er mwyn sicrhau lefelau priodol o gydlyniant â dulliau o reoleiddio deallusrwydd artiffisial sy'n datblygu, byddwn yn parhau i weithio gyda phartneriaid rhyngwladol ar y gallu i ryngweithredu wrth reoleiddio, gan gynnwys safonau technegol a thechnegau sicrwydd, er mwyn ei gwneud yn haws i gwmnïau yn y DU i ddenu buddsoddiad tramor a masnachu'n rhyngwladol. Am ragor o fanylion, gweler adran 5.3 a'n hymateb i gwestiynau ynglŷn ag adnoddau ar gyfer deallusrwydd artiffisial dibynadwy yn 6.6.

124. Ochr yn ochr â hyn, rydym wedi cyhoeddi gwasanaeth rheoleiddio peilot newydd i'w gynnal gan y Fforwm Cydweithredu Rheoleiddio Digidol i'w gwneud yn haws i arloeswyr deallusrwydd artiffisial a digidol ddeall a symud drwy'r dirwedd reoleiddio (gweler ein hymateb i gwestiynau ynglŷn â phyllau tywod deallusrwydd artiffisial am ragor o fanylion: adran 6.10).

125. Rydym yn ymrwymedig o hyd i'r dull iteraidd a nodwyd yn y papur gwyn, a disgwyliwn y bydd angen i'n fframwaith ddatblygu wrth i risgiau newydd neu fylchau mewn rheoleiddio ddod i'r amlwg. Bydd ein gweithgareddau monitro a gwerthuso yn asesu a ddylem wneud newidiadau i'n fframwaith, pryd, a sut, gan gasglu tystiolaeth o amrywiaeth eang o ffynonellau. Rhown ragor o fanylion yn ein hymateb i gwestiynau ynglŷn â monitro a gwerthuso yn adran 6.4.

126. Mae'n galonogol nodi bod ymatebwyr wedi cymeradwyo amrywiaeth eang o sefydliadau yn y DU fel partneriaid defnyddiol i gyflawni'r gweithgareddau canolog arfaethedig. Fel y dywedwyd gennym yn y papur gwyn, y llywodraeth fydd yn cyflawni'r swyddogaeth ganolog i ddechrau, gan weithio mewn partneriaeth â rheoleiddwyr a chyfranogwyr allweddol yn yr ecosystem deallusrwydd artiffisial. Prif rôl y llywodraeth fydd manteisio ar y gweithgareddau sy'n bodoli eisoes lle y bo modd a sicrhau bod yr holl weithgareddau angenrheidiol i hyrwyddo arloesi'n gyfrifol ym maes deallusrwydd artiffisial yn mynd rhagddynt.

## 6.4. Monitro a gwerthuso'r fframwaith

**15. A ydych chi'n cytuno â'n dull gweithredu cyffredinol o fonitro a gwerthuso?**

**16. Beth yw'r ffordd orau o fesur effaith ein fframwaith?**

**17. A ydych chi'n cytuno bod ein dull gweithredu yn taro'r cydbwysedd cywir rhwng cefnogi arloesedd ym maes deallusrwydd artiffisial; ymdrin â risgiau hysbys, wedi'u blaenoriaethu; a diogelu'r fframwaith rheoleiddio deallusrwydd artiffisial rhag datblygiadau yn y dyfodol?**

Crynodeb o gwestiynau 15-17:

127. Roedd y mwyafrif o'r ymatebwyr yn cytuno â'r dull cyffredinol o fonitro a gwerthuso, gan ganmol y ddolen adborth arfaethedig â diwydiant a chymdeithas sifil fel modd i ddeall pa mor effeithiol yw'r fframwaith.

128. Pwysleisiodd ychydig dros chwarter yr ymatebwyr mai drwy ymgysylltu ag amrywiaeth eang o randdeiliaid y byddai modd meithrin y ddealltwriaeth fwyaf gwerthfawr. Dadleuodd llawer o blaid cynnwys cymdeithas sifil ehangach a chynrychiolwyr defnyddwyr er mwyn sicrhau bod lleisiau y tu allan i'r diwydiant technoleg yn cael eu clywed, yn ogystal ag ymgysylltu rheolaidd ag arbenigwyr diwydiant ac ymchwil. Pwysleisiodd ymatebwyr hefyd y byddai ymgysylltu rhyngwladol yn allweddol i gysoni dulliau gweithredu rhwng awdurdodaethau yn effeithiol.

129. Roedd ymatebwyr am weld rhagor o fanylion am ymarferoldeb y fframwaith monitro a gwerthuso, gan gynnwys sut y caiff data eu casglu a'u defnyddio i fesur llwyddiant. Awgrymodd bron traean o'r ymatebwyr y dylai effaith y fframwaith gael ei mesur drwy amrywiaeth o ffynonellau data, gan argymhell y dylid casglu data ar ddangosyddion allweddol yn ogystal â defnyddio asesiadau o'r effaith.

130. Roedd hanner yr ymatebwyr yn cytuno bod y dull gweithredu, i bob golwg, yn taro'r cydbwysedd cywir rhwng cefnogi arloesi mewn deallusrwydd artiffisial; ymdrin â risgiau hysbys, wedi'u blaenoriaethu; a diogelu'r fframwaith rheoleiddio deallusrwydd artiffisial rhag datblygiadau yn y dyfodol. Fodd bynnag, roedd rhai o'r ymatebwyr yn anghytuno, gan ddadlau bod y dull gweithredu yn rhoi mwy o flaenoriaeth i arloesi ym maes deallusrwydd artiffisial a thwf economaidd na diogelwch a lliniaru risgiau sy'n gysylltiedig â deallusrwydd artiffisial.

**Ymateb:**

131. Mae'n bleser gennym nodi'r adborth cadarnhaol ynglŷn â'n dull arfaethedig o fonitro a gwerthuso'r fframwaith. Drwy weithgareddau monitro a gwerthuso, bydd modd i ni adolygu'r ffordd y caiff y fframwaith rheoleiddio deallusrwydd artiffisial ei roi ar waith ym mhob rhan o'r economi ac maent wrth wraidd ein dull gweithredu iteradd. Bydd yn sicrhau bod y gyfundrefn yn gweithio yn ôl y bwriad: gan ymateb yn weithredol i risgiau wedi'u blaenoriaethu, cefnogi arloesi, a sicrhau manteision mwyaf deallusrwydd artiffisial ledled y DU. Rydym yn cytuno â'r ymatebwyr, wrth i ni roi'r fframwaith a nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial ar waith, y bydd monitro a gwerthuso yn ei gwneud yn bosibl i'r llywodraeth adnabod problemau posibl ac addasu'r fframwaith mewn ymateb i hynny os bydd angen.

132. Rydym yn cydnabod y pryderon cynyddol y byddwn yn wynebu mwy o risgiau i ddiogelwch sy'n gysylltiedig â deallusrwydd artiffisial o bosibl wrth i'r technolegau hyn gael eu defnyddio'n fwyfwy. Rydym yn cydnabod bod llawer o'r pryderon hyn yn ymwneud â galluoedd uwch y systemau deallusrwydd artiffisial mwy galluog. Dyna pam rydym yn ymrwymedig o hyd i ddull hyblyg a fydd yn datblygu wrth i risgiau newydd neu fylchau mewn rheoleiddio ddod i'r amlwg. Cyflwynir ein syniadaeth gychwynol ynglŷn â mesurau newydd posibl sydd wedi'u targedu at ddatblygwyr modelau deallusrwydd artiffisial diben cyffredinol tra galluog yn adran 5.2. Bydd y Sefydliad Diogelwch Deallusrwydd Artiffisial yn hyrwyddo galluoedd diogelwch ym maes deallusrwydd artiffisial er budd y cyhoedd, gan alluogi'r llywodraeth i ymateb i ddatblygiadau technolegol arloesol. Bydd ein gweithgarwch monitro a gwerthuso yn adeiladu ar waith gan y Sefydliad, ein hasesiad risg traws-sector, ac adborth gan randdeiliaid er mwyn deall sut mae'r fframwaith rheoleiddio yn perfformio. Bydd ein gwerthusiad yn ystyried a yw'r fframwaith yn cyflawni'r amcanion a nodwyd yn y papur gwyn yn effeithiol, gan gynnwys ennyn ymddiriedaeth y cyhoedd drwy ymdrin â risgiau posibl mewn ffordd briodol.

133. Nodwn y pwyslais gan ymatebwyr ar ddefnyddio'r data, y metrigau, a'r ffynonellau cywir i werthuso pa mor dda y mae'r fframwaith rheoleiddio yn cyflawni. Rydym yn cytuno ei bod yn allweddol i effeithiolrwydd y fframwaith ein bod yn cael y mesurau o lwyddiant yn gywir, ac rydym wrthi'n gweithio ar hyn wrth i ni ddatblygu ein fframwaith monitro a gwerthuso i'w gyhoeddi. Byddwn yn cynnal ymgynghoriad wedi'i dargedu ar ein cynllun arfaethedig i asesu'r fframwaith gyda nifer o randdeiliaid yn y gwanwyn. Fel rhan o'r gwaith hwn, byddwn yn ceisio barn fanwl ar ein metrigau a'n ffynonellau data arfaethedig.

## 6.5. Galluoedd rheoleiddwyr

**18. A ydych chi'n cytuno mai rheoleiddwyr sydd yn y sefyllfa orau i gymhwyso'r egwyddorion ac mai'r llywodraeth sydd yn y sefyllfa orau i oruchwylio a chyflawni swyddogaethau canolog?**

**19. Fel rheoleiddiwr, pa gymorth fyddai ei angen arnoch i gymhwyso'r egwyddorion mewn ffordd gymesur sydd o blaid arloesi?**

**20. A ydych chi'n cytuno mai tîm cyfunol o arbenigwyr ar ddeallusrwydd artiffisial fyddai'r ffordd fwyaf effeithiol o lenwi bylchau mewn galluoedd a helpu rheoleiddwyr i gymhwyso'r egwyddorion?**

Crynodeb o gwestiynau 18-20:

134. Roedd bron pob un o'r ymatebwyr yn cytuno mai rheoleiddwyr sydd yn y sefyllfa orau i arwain y gwaith o roi'r egwyddorion ar waith, ac mai'r llywodraeth sydd yn y sefyllfa orau i oruchwylio a chyflawni'r swyddogaethau canolog. Fodd bynnag, roedd ymatebwyr yn dadlau y byddai angen i'r llywodraeth wella galluoedd rheoleiddwyr er mwyn i'r dull gweithredu hwn fod yn effeithiol. Roedd rhai ymatebwyr yn pryderu ynglŷn â diffyg corff penodol i helpu i roi'r fframwaith arfaethedig ar waith a'i oruchwylio, gyda rhai yn gofyn am ddeddfwriaeth ynglŷn â deallusrwydd artiffisial a rheoleiddiwr newydd ar gyfer deallusrwydd artiffisial.

135. Er bod y rheoleiddwyr yn gefnogol i'r dull gweithredu arfaethedig ar y cyfan, awgrymodd dros chwarter y rhai a ymatebodd i Gwestiwn 19 y byddai mwy o arbenigedd mewn deallusrwydd artiffisial yn eu helpu i gymhwyso'r egwyddorion o fewn eu cylchoedd gwaith presennol. Ar y cyfan, nododd y rheoleiddwyr wahanol lefelau o arbenigedd

technegol a galluoedd deallusrwydd artiffisial. Roedd rhai yn teimlo y byddai mwy o allu sefydliadol ac adnoddau ychwanegol yn eu helpu i ymgymryd â chyfrifoldebau newydd mewn perthynas â deallusrwydd artiffisial a deall ble a sut y caiff ei ddefnyddio yn eu priod feysydd.

136. Nododd y rheoleiddwyr hefyd fod deallusrwydd artiffisial yn achosi rhai heriau o ran cydgyssylltu ar draws eu priod feysydd a sectorau, ac nad oedd rhai risgiau a oedd yn dod i'r amlwg mewn perthynas â deallusrwydd artiffisial yn perthyn yn glir i unrhyw gylch gwaith presennol penodol. Pwysleisiodd ychydig dros chwarter y rheoleiddwyr a ymatebodd i Gwestiwn 19 y byddai cydweithio agos rhwng rheoleiddwyr a'r swyddogaethau canolog arfaethedig yn helpu i greu gofynion sector penodol ystyrlon ac atal dyblygu gwaith.

137. Roedd y mwyafrif o'r ymatebwyr yn cytuno mai tîm cyfunol o arbenigwyr ar ddeallusrwydd artiffisial fyddai'r ffordd fwyaf effeithiol o ymdrin â'r lefelau gwahanol o allu yn y dirwedd reoleiddio. Dadleuodd ymatebwyr o blaid cronfa amrywiol ac amlddisgyblaethol er mwyn dwyn ynghyd arbenigwyr deallusrwydd artiffisial technegol sydd â gwybodaeth reoleiddio am sectorau penodol, arbenigwyr diwydiant, a chymdeithas sifil. Dadleuodd ymatebwyr y byddai hyn yn sicrhau bod rheoleiddwyr yn ystyried amrywiaeth eang o safbwyntiau wrth gymhwyso'r egwyddorion deallusrwydd artiffisial traws-sector.

#### Ymateb:

138. Mae'n galonogol bod ymatebwyr yn cytuno'n fras â'r dull gweithredu arfaethedig a arweinir gan reoleiddwyr er mwyn rhoi'r egwyddorion ar waith, gyda'r llywodraeth yn goruchwylio ac yn cyflawni'r swyddogaeth ganolog. Fel yr amlinellwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, ein rheoleiddwyr arbenigol presennol sydd yn y sefyllfa orau i ymgymryd â dadansoddiad manwl o'r risgiau a gweithgareddau gorfodi yn eu meysydd o arbenigedd. Byddwn yn parhau i weithio'n agos gyda'r rheoleiddwyr er mwyn sicrhau yr ymdrinnir yn ddigonol â risgiau posibl a achosir gan ddeallusrwydd artiffisial o dan reolaeth y gyfraith. Yn unol â'n dull gweithredu iteraid, byddwn yn ceisio addasu'r fframwaith, gan gynnwys y bensaerniaeth reoleiddio, os bydd gwaith dadansoddi yn dangos bod hyn yn angenrheidiol ac yn effeithiol.

139. Fel y nododd ymatebwyr i'r ymgynghoriad, er mwyn rheoleiddio deallusrwydd artiffisial yn effeithiol bydd yn rhaid i'n rheoleiddwyr feddu ar y sgiliau, yr adnoddau, a'r arbenigedd cywir. Er mwyn helpu rheoleiddwyr i allu ymaddasu ac ymateb i'r risgiau a'r cyfleoedd sy'n codi o ganlyniad i ddeallusrwydd artiffisial yn eu priod feysydd, rydym heddiw yn cyhoeddi buddsoddiad o £10 miliwn ar gyfer uwchsgilio technegol. Byddwn yn gweithio'n agos gyda rheoleiddwyr i nodi'r cyfleoedd mwyaf addawol i ddefnyddio'r cyllid hwn, gan gynnwys llunio model cyflawni sy'n gallu cyflawni'r amcanion bwriadedig yn fwy effeithiol na'r gronfa ganolog o arbenigedd a gynigir yn y papur gwyn ar reoleiddio deallusrwydd artiffisial. Yn benodol, mae adborth rheoleiddwyr wedi dangos bod angen i ni eu helpu i ddatblygu adnoddau a sgiliau yn eu priod feysydd – ond drwy gydweithio lle y bo'n briodol – a chynnig y cymorth sy'n gyson â'u hannibyniaeth ac sy'n ei chefnogi. Gan fod galluoedd ac adnoddau yn amrywio rhwng rheoleiddwyr, ein bwriad yw y bydd y gronfa hon yn enwedig yn galluogi'r rheoleiddwyr hynny y mae eu harbenigedd mewn deallusrwydd artiffisial yn llai aeddfed i wneud gwaith ymchwil a datguddio dirnadaethau sylfaenol i ddatblygu neu addasu adnoddau ymarferol er mwyn sicrhau cydymffurfiaeth ar gyfer dyfodol a hwylusir gan ddeallusrwydd artiffisial.

140. At hynny, fel y nodwyd yn yr ymateb i adolygiad trawsbynciol yr Athro y Fonesig Angela McLean o reoleiddio technolegau mewn ffordd sydd o blaid arloesi,<sup>110</sup> mae'r llywodraeth hefyd yn ystyried sut i gefnogi rheoleiddwyr ymhellach er mwyn iddynt ddatblygu'r sgiliau arbenigol sydd eu hangen i reoleiddio technolegau sy'n datblygu, gan gynnwys mwy o hyblygrwydd ynglŷn â chyflogau ac amodau. Mae hyn yn adeiladu ar gynlluniau sydd eisoes ar waith i gefnogi secondiadau rhwng adrannau'r llywodraeth, rheoleiddwyr, y byd academiaidd, a diwydiant.

141. Rydym yn cydnabod pryderon rheoleiddwyr y gall deallusrwydd artiffisial achosi heriau o ran cydgysylltu. Yn y papur gwyn, cynigiwyd nifer o weithgareddau canolog i gefnogi rheoleiddwyr a sicrhau bod y dirwedd reoleiddio ar gyfer deallusrwydd artiffisial yn gyson ac yn gydlynol. Er mwyn hwyluso cydweithrediad trawsbynciol a sicrhau bod y fframwaith rheoleiddio cyffredinol yn gweithredu yn ôl y bwriad, rydym yn datblygu ein gweithgareddau cydgysylltu rheoleiddiol. Bydd y gweithgareddau cydgysylltu hyn yn rhan o'n swyddogaeth ganolog yn y llywodraeth ochr yn ochr â'n gweithgareddau asesu risg deallusrwydd artiffisial (ceir rhagor o fanylion yn adran 5.1.2). Er mwyn cefnogi dull gweithredu cydlynol rhwng sectorau, rydym hefyd yn cyhoeddi canllawiau cychwynnol i reoleiddwyr, ochr yn ochr â'r ymateb hwn ynglŷn â sut y dylid cymhwyso'r egwyddorion deallusrwydd artiffisial traws-sector o fewn eu cylchoedd gwaith presennol.

142. Nodwn bwyslais ymatebwyr ar dryloywder a'r angen i ddiwydiant a chymdeithas sifil fod yn amlwg yn y fframwaith rheoleiddio deallusrwydd artiffisial. Rydym yn cytuno y bydd sefydlu dolenni adborth â diwydiant, y byd academiaidd a chymdeithas sifil yn allweddol i fesur effeithiolrwydd y fframwaith. Bydd ein swyddogaeth ganolog yn ymgysylltu â rhanddeiliaid er mwyn sicrhau bod amrywiaeth eang o leisiau yn cael eu clywed a'u hystyried: gan roi eglurder, meithrin ymddiriedaeth, sicrhau'r gallu i ryngweithredu, a hysbysu'r llywodraeth pan fo angen addasu'r fframwaith.

## 6.6. Adnoddau ar gyfer deallusrwydd artiffisial dibynadwy

**21. Pa adnoddau anrheoleiddiol ar gyfer deallusrwydd artiffisial dibynadwy fyddai fwyaf defnyddiol i helpu sefydliadau i ymgorffori egwyddorion rheoleiddio deallusrwydd artiffisial mewn prosesau busnes sy'n bodoli eisoes?**

Crynodeb o gwestiwn 21:

143. Cafwyd cefnogaeth gref i ddefnyddio safonau technegol a thechnegau sicrwydd, gyda rhai ymatebwyr yn cytuno y byddai'r naill a'r llall yn helpu sefydliadau i ymgorffori'r egwyddorion deallusrwydd artiffisial mewn prosesau busnes sy'n bodoli eisoes. Gwnaeth llawer o'r ymatebwyr ganmol gwaith Hwb Safonau Deallusrwydd Artiffisial y DU a'r Ganolfan Moeseg Data ac Arloesi ar sicrwydd deallusrwydd artiffisial. Er bod rhai ymatebwyr wedi nodi y byddai llai o faich cydymffurfio ar fusnesau pe bai adnoddau a phrosesau yn gyson rhwng sectorau, nododd eraill bwysigrwydd adnoddau a phrosesau ychwanegol ar gyfer sectorau penodol. Awgrymodd ymatebwyr hefyd y dylid ategu safonau technegol ag astudiaethau achos ac enghreifftiau o arferion da.

144. Dadleuodd ymatebwyr y byddai adnoddau a thechnegau safonol ar gyfer nodi a lliniaru risgiau posibl sy'n gysylltiedig â deallusrwydd artiffisial hefyd yn helpu sefydliadau i ymgorffori'r egwyddorion deallusrwydd artiffisial. Nododd rhai fod technegau sicrwydd

<sup>110</sup> [Response to Professor Dame Angela McLean's Pro-Innovation Regulation of Technologies Review: Cross Cutting](#), Trysorlys EF, 2023.

megis asesiadau o'r effaith ac asesiadau risg, monitro perfformiad modelau, gwerthuso ansicrwydd modelau, ac ymarferion tîm coch yn ddefnyddiol iawn i nodi risgiau deallusrwydd artiffisial. Argymhellodd ychydig o'r ymatebwyr dechnegau sicrwydd y gellir eu defnyddio i ganfod ac atal problemau megis gadael i bethau lithro o ran lliniaru risgiau sy'n ymwneud â data. Er eu bod yn cymeradwyo rôl adnoddau ar gyfer deallusrwydd artiffisial dibynadwy, nododd nifer bach o'r ymatebwyr eu bod yn awyddus i weld mesurau rheoleiddio llymach, megis gofynion statudol ar gyfer cymwysiadau deallusrwydd artiffisial risg uchel neu gorff gwarchod ar gyfer modelau sylfaen.

145. Roedd ymatebwyr yn teimlo y gall adnoddau a thechnegau megis metrigau tegwch, adroddiadau tryloywder, a chanllawiau moeseg deallusrwydd artiffisial sefydliadol gefnogi'r defnydd cyfrifol o ddeallusrwydd artiffisial a hefyd feithrin ymddiriedaeth y cyhoedd yn y dechnoleg ar yr un pryd. Nododd ymatebwyr eu bod yn awyddus i fodolau deallusrwydd artiffisial gael eu dilysu gan drydydd parti drwy archwiliadau o duedd, cynlluniau labelu i ddefnyddwyr, ac ardystio allanol yn erbyn safonau technegol.

146. Nododd ychydig o'r ymatebwyr fanteision i sefydliadau a defnyddwyr fel ei gilydd pe bai dulliau o lywodraethau deallusrwydd artiffisial yn cael eu cysoni'n rhyngwladol. Roedd rhai yn cymeradwyo safonau technegol ynglŷn â'r gallu i ryngweithredu ar gyfer deallusrwydd artiffisial, gan ganmol sefydliadau datblygu safonau byd-eang megis y Sefydliad Safoni Rhyngwladol (ISO) a Sefydliad y Peirianwyr Trydanol ac Electroneg (IEEE). Nododd eraill gryfder ystod o waith rhyngwladol ar ddeallusrwydd artiffisial, gan gynnwys gwaith gan wledydd unigol, megis Fframwaith Rheoli Risgiau Deallusrwydd Artiffisial Sefydliad Safonau a Thechnoleg Cenedlaethol (NIST) UDA a Sefydliad Dilysu Deallusrwydd Artiffisial Singapore, ochr yn ochr â gwaith ar lywodraethu rhyngwladol gan gyrff amlochrog megis y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, y Cenhedloedd Unedig, a fforwm y G7.

## Ymateb:

147. Mae'n bleser gennym weld cymaint o gefnogaeth gref o blaid parhau i ddatblygu a mabwysiadu safonau technegol a thechnegau sicrwydd ar gyfer deallusrwydd artiffisial. Bydd yr adnoddau hyn yn helpu sefydliadau i roi ein hegwyddorion rheoleiddio arfaethedig ar waith, arloesi mewn ffordd gyfrifol, ac ennyn hyder y cyhoedd. Rydym yn cydnabod, mewn rhai achosion, y bydd yn bwysig cael technegau sicrwydd a safonau technegol sy'n benodol i gyd-destun, cymhwysiad, neu sector arbennig. Dyna pam y gwnaethom nodi dull haenog o gyflwyno safonau technegol yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, sy'n annog rheoleiddwyr i adeiladu ar adnoddau sector-amheuol y gellir eu defnyddio'n eang lle y bo'n briodol.<sup>111</sup>

148. Rydym yn croesawu'r ganmoliaeth am Hwb Safonau Deallusrwydd Artiffisial y DU a'r Ganolfan Moeseg Data ac Arloesi. Mae'r Hwb, a lansiwyd ym mis Hydref 2022, yn dwyn ynghyd arbenigedd technegol y DU ar safonau deallusrwydd artiffisial, gan gynnwys Sefydliad Alan Turing, y Sefydliad Safonau Prydeinig, a'r Labordy Ffisegol Cenedlaethol, i roi hyfforddiant a gwybodaeth am y dirwedd safonau deallusrwydd artiffisial ryngwladol, gymhleth. Cyhoeddodd y Ganolfan Moeseg Data ac Arloesi Bortffolio o Dechnegau Sicrwydd Deallusrwydd Artiffisial ym mis Mehefin 2023, a oedd yn cynnwys enghreifftiau go iawn er mwyn helpu i ddatblygu deallusrwydd artiffisial dibynadwy, y nododd ymatebwyr y byddai'n ddefnyddiol.<sup>112</sup> Mae'r Portffolio hefyd yn rhan o'r Catalog Adnoddau a Metrigau ar gyfer Deallusrwydd Artiffisial Dibynadwy y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, sy'n rhannu astudiaethau achos y Ganolfan i gynulleidfa ryngwladol

<sup>111</sup> [AI regulation: a pro-innovation approach](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>112</sup> [CDEI portfolio of AI assurance techniques](#), y Ganolfan Moeseg Data ac Arloesi; Yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

Hefyd, lansiodd y Ganolfan “Her Arloesi Tegwch” ym mis Hydref i helpu i ddatblygu atebion technolegol-gymdeithasol newydd i ymdrin â thuedd a gwahaniaethu mewn systemau deallusrwydd artiffisial.<sup>113</sup> Heddiw, rydym yn cyhoeddi bod y Ganolfan Moeseg Data ac Arloesi yn newid ei henw i'r Uned Mabwysiadu Technoleg yn Gyfrifol er mwyn adlewyrchu'n fwy cywir ei rôl yn yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg i ddatblygu adnoddau a thechnegau sy'n ei gwneud yn bosibl i ddeallusrwydd artiffisial gael ei fabwysiadu mewn ffordd gyfrifol yn y sector preifat a'r sector cyhoeddus. Eleni, bydd yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg yn cyhoeddi “Introduction to AI assurance” er mwyn hyrwyddo gwerth sicrwydd deallusrwydd artiffisial ymhellach.

149. Rydym yn nodi yr hoffai ymatebwyr weld mwy o adnoddau a thechnegau safonol i nodi a rheoli risgiau deallusrwydd artiffisial. Cyn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial ym mis Tachwedd 2023, cyhoeddwyd “Emerging processes for frontier AI safety” gennym er mwyn helpu i ysgogi trafodaeth ynglŷn â sut beth fyddai prosesau diogelwch da ar gyfer systemau deallusrwydd artiffisial uwch.<sup>114</sup> Mae'r ddogfen yn rhoi ciplun o syniadau addawol, prosesau sy'n datblygu, ac arferion cysylltiedig o ran diogelwch deallusrwydd artiffisial. Bwriedir i'r ddogfen fod yn gyfeirbwynt i lywio'r broses o ddatblygu polisiau diogelwch sefydliadau deallusrwydd artiffisial ffin eithaf yn ogystal â bod yn arweinyfr i'r rhai sy'n darllen y polisiau hyn. Mae'n amlinellu syniadau cynnar ynglŷn ag arferion arloesi mewn datblygiadau deallusrwydd artiffisial ffin eithaf, gan gynnwys gwerthuso modelau ac ymarferion tîm coch, meithrin gallu ar raddfa gyfrifol, a phrosesau adrodd a rhannu gwybodaeth am fodelau. Yn 2024, byddwn yn annog cwmnïau deallusrwydd artiffisial i ddatblygu eu polisiau o ran diogelwch deallusrwydd artiffisial a meithrin gallu ar raddfa gyfrifol. Fel rhan o'r gwaith hwn, byddwn yn diweddarau ein canllaw ar brosesau sy'n datblygu erbyn diwedd y flwyddyn. Yn fwy cyffredinol, nodwn ddatblygiad safonau technolegol byd-eang perthnasol sy'n rhoi arweiniad ar reoli risg sy'n gysylltiedig â deallusrwydd artiffisial. Er enghraifft, bydd safon ISO 42001 yn helpu sefydliadau i reoli eu systemau deallusrwydd artiffisial mewn ffordd sy'n ennyn ymddiriedaeth.

150. Yn y papur gwyn, nodwn fod ymateb i risgiau ac ennyn ymddiriedaeth y cyhoedd yn ysgogwyr rheoleiddio allweddol. Felly, rydym yn deall pwyslais ymatebwyr ar adnoddau i ennyn ymddiriedaeth y cyhoedd fel ffordd allweddol o sicrhau arloesi'n gyfrifol ym maes deallusrwydd artiffisial. Mae'r Uned Mabwysiadu Technoleg yn Gyfrifol (y Ganolfan Moeseg Data ac Arloesi gynt) yn cynnwys tîm Deall y Cyhoedd arbenigol sy'n ymgysylltu â'r cyhoedd a chymunedau yr effeithir arnynt yn rheolaidd er mwyn meithrin dealltwriaeth ddofn o agweddau'r cyhoedd tuag at ddeallusrwydd artiffisial.<sup>115</sup> Defnyddir y ddealltwriaeth hon gan yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg a'r llywodraeth yn fwy cyffredinol i gysoni ein dulliau rheoleiddiol o ymdrin â deallusrwydd artiffisial â gwerthoedd cyhoeddus ac ennyn ymddiriedaeth yn y technolegau hyn. Mae'r Adran dros Wyddoniaeth, Arloesedd a Thechnoleg a'r Swyddfa Digidol a Data Ganolog wedi datblygu'r Safon Cofnodi Tryloywder Algorithmig hefyd er mwyn helpu sefydliadau yn y sector cyhoeddus i ddarparu gwybodaeth glir am yr adnoddau algorithmig y maent yn eu defnyddio i gefnogi penderfyniadau.<sup>116</sup> Yn sgil treialu'r safon yn llwyddiannus, a chyhoeddi fersiwn drawslywodraethol gymeradwy y llynedd, byddwn bellach yn ei gwneud yn ofynnol i bob un o adrannau'r llywodraeth ddefnyddio'r safon a bwriadwn ymestyn hyn i'r sector cyhoeddus yn fwy cyffredinol dros amser.

<sup>113</sup> [Fairness Innovation Challenge](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg; Innovate UK, 2023.

<sup>114</sup> [Emerging Processes for Frontier AI Safety](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>115</sup> Am drosolwg o ymchwil ddiweddarach yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg i agweddau'r cyhoedd tuag at ddata a deallusrwydd artiffisial, gweler: [Public attitudes to data and AI: Tracker survey \(Wave 3\)](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

<sup>116</sup> Y Safon Cofnodi Tryloywder Algorithmig neu ATRS yn Saesneg. Am ragor o fanylion gweler adran 5.1.



151. Rydym yn cytuno ag ymatebwyr y bydd cydweithrediad rhyngwladol ynglŷn â llywodraethu deallusrwydd artiffisial yn allweddol i liniaru risgiau sy'n gysylltiedig â deallusrwydd artiffisial yn llwyddiannus ac ennyn ymddiriedaeth y cyhoedd mewn deallusrwydd artiffisial. Yn yr Uwchgyhadledd gyntaf erioed ar Ddiogelwch Deallusrwydd Artiffisial, cynulliwyd grŵp o gynrychiolwyr o bedwar ban byd i bennu llwybr newydd ar gyfer gweithredu rhyngwladol ar y cyd i ddeall cyfleoedd a risgiau deallusrwydd artiffisial ffin eithaf. Rydym hefyd yn parhau i gydweithio'n rhyngwladol ar lywodraethu deallusrwydd artiffisial, yn ddwyochrog a thrwy sawl fforwm amlochrog. Er enghraifft, mae'r DU yn chwarae rhan bwysig mewn trafodaethau ynglŷn â deallusrwydd artiffisial yn y CU, Cyngor Ewrop, y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, fforwm y G7, y Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial a fforwm y G20. Dylid nodi'n benodol fod y DU wedi gweithio'n agos gyda phartneriaid G7 i gyd-drafod y Codau Ymddygiad a'r Egwyddorion Arweiniol ynglŷn â datblygu systemau deallusrwydd artiffisial uwch, fel rhan o Broses Deallusrwydd Artiffisial Hiroshima. Mae'r DU yn gwbl gefnogol bod polisi a safonau technegol deallusrwydd artiffisial yn cael eu datblygu mewn ffordd agored a chynhwysol yn fyd-eang, sy'n cynnwys rhanddeiliaid ac sy'n seiliedig ar gonsensws. Rydym yn cefnogi ymwneud rhanddeiliaid yn y DU â Sefydliadau Datblygu Safonau er mwyn sicrhau manteision safonau technegol byd-eang yma yn y DU a chyflwyno safonau technegol digidol byd-eang sydd wedi'u ffurfio gan werthoedd democrataidd.

## 6.7. Syniadau i gloi

**22. A oes gennych unrhyw syniadau eraill ynglŷn â'n dull gweithredu cyffredinol? Nodwch unrhyw gyfleoedd coll, diffygion, a bylchau yn ein fframwaith.**

### Crynodeb o gwestiwn 22:

152. Roedd rhai ymatebwyr yn teimlo y byddai'n fuddiol pe bai'r fframwaith rheoleiddio deallusrwydd artiffisial a nodwyd yn y papur gwyn yn cynnwys canllawiau manylach ar risgiau sy'n gysylltiedig â deallusrwydd artiffisial. Roedd rhai yn awyddus i weld mesurau llymach ar gyfer risgiau difrifol, yn enwedig o ran y defnydd o ddeallusrwydd artiffisial mewn cyd-destun lle mae diogelwch yn hanfodol. Awgrymodd ymatebwyr y byddai'r fframwaith yn gliriach pe bai'r llywodraeth yn rhoi categorïau o risg ar gyfer defnyddiau penodol o ddeallusrwydd artiffisial megis gorfodi'r gyfraith a gweithleoedd. Pwysleisiodd ymatebwyr eraill y gall deallusrwydd artiffisial achosi neu gyflymu risgiau sylweddol ynglŷn â phreifatrwydd a thorri cyfraith diogelu data, seiberymosodiadau, ymyrryd mewn etholiadau, camwybodaeth, ymyrryd â hawliau dynol, cynaliadwyedd amgylcheddol, a materion sy'n ymwneud â chystadleuaeth. Roedd ychydig o'r ymatebwyr yn pryderu am y risg ddifodol bosibl sy'n codi o ganlyniad i ddeallusrwydd artiffisial. Roedd llawer o'r ymatebwyr yn teimlo bod technolegau deallusrwydd artiffisial yn datblygu'n gynt na'r prosesau rheoleiddio.

153. Roedd rhai ymatebwyr yn dadlau bod llwyddiant y fframwaith yn dibynnu ar ddigon o gydgyssylltu rhwng rheoleiddwyr er mwyn cynnig dull clir a chyson o ymdrin â deallusrwydd artiffisial rhwng sectorau a marchnadoedd. Nododd ymatebwyr hefyd fod sectorau gwahanol yn wynebu manteision a risgiau penodol mewn cysylltiad â deallusrwydd artiffisial, gan awgrymu y byddai angen i'r fframwaith gydbwyso'r cysondeb a gynigir drwy ofynion traws-sector â chywirdeb dulliau gweithredu ar gyfer sectorau penodol. Yn benodol, nododd ymatebwyr y dylai unrhyw reolau neu gyrff newydd i reoleiddio deallusrwydd

artiffisial gael eu hadeiladu ar gylchoedd gwaith statudol presennol rheoleiddwyr a safonau rheoleiddio perthnasol. Nododd ymatebwyr hefyd y byddai angen adnoddau digonol ar reoleiddwyr ar ffurf arbenigedd a sgiliau technegol i roi'r fframwaith ar waith yn effeithiol.

154. Pwysleisiodd ymatebwyr yn rheolaidd bwysigrwydd cysoni rhyngwladol i reoleiddio deallusrwydd artiffisial yn effeithiol. Awgrymodd rhai ymatebwyr y dylai'r DU weithio tuag at greu ecosystem reoleiddiol ar gyfer deallusrwydd artiffisial sy'n gyson yn rhyngwladol drwy ddatblygu fframwaith safon aur a hyrwyddo arferion gorau drwy sianeli amlochrog allweddol megis y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, y CU, y Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial, fforwm y G7, fforwm y G20, a Chyngor Ewrop. Nododd ymatebwyr y byddai dulliau o reoleiddio deallusrwydd artiffisial sy'n anghyson neu sy'n gorgyffwrdd yn achosi beichiau cydymffurfio sylweddol. Roedd ymatebwyr yn dadlau y gall cydweithredu rhyngwladol gefnogi arloesi'n gyfrifol ym maes deallusrwydd artiffisial yn y DU drwy greu rheolau clir a phendant sy'n ei gwneud yn bosibl i fuddsoddiadau symud ar draws sawl marchnad. Awgrymodd ymatebwyr hefyd y dylid sefydlu gweithgorau dwyochrog â phartneriaid strategol allweddol er mwyn rhannu arbenigedd. Pwysleisiodd rhai ymatebwyr y dylai dull gweithredu'r DU o blaid arloesi gael ei gyflwyno'n ddi-oed er mwyn aros yn gystadleuol mewn tirwedd ryngwladol sy'n newid yn gyflym.

## Ymateb:

155. Rydym yn cydnabod yr hoffai llawer o'r ymatebwyr weld mwy o fanylion ynglŷn â'r broses o roi'r fframwaith ar waith a nodwyd yn y papur gwyn, yn enwedig o ran risgiau sy'n gysylltiedig â deallusrwydd artiffisial. Rydym eisoes wedi dechrau cyflwyno'r cynigion a nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, gan weithio'n gyflym i bennu gweithgareddau asesu risg traws-economi canolog o fewn y llywodraeth er mwyn nodi, mesur, a lliniaru risgiau. Gan adeiladu ar y gwaith hwn, gwnaethom gyhoeddi ymchwil i alluoedd a risgiau deallusrwydd artiffisial ffin eithaf i'w thrafod yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial.<sup>117</sup> Amlinellodd dystiolaeth gychwynnol ynglŷn â'r systemau deallusrwydd artiffisial lefel uchaf a sut y gall eu galluoedd a'u risgiau barhau i ddatblygu. Mae ansicrwydd mawr yn y dystiolaeth yn awgrymu bod angen gwneud rhagor o ymchwil.

156. Eleni, byddwn yn ymgynghori ar gofrestr risgiau deallusrwydd artiffisial draws-economi, gan geisio barn arbenigol ar ein methodoleg o asesu risg ac a ydym wedi cofnodi risgiau sy'n gysylltiedig â deallusrwydd artiffisial mewn ffordd gynhwysfawr. Bydd y Sefydliad Diogelwch Deallusrwydd Artiffisial yn datblygu gwybodaeth y byd am ddiogelwch deallusrwydd artiffisial drwy archwilio, gwerthuso a phrofi systemau deallusrwydd artiffisial uwch yn ofalus. Bydd yn gwneud ymchwil sylfaenol ynglŷn â sut i gadw pobl yn ddiogel yn wyneb cynnydd technolegol cyflym ac anrhagweladwy.

157. Yn y papur gwyn, cynigiwyd dull hyblyg o reoleiddio deallusrwydd artiffisial, yn seiliedig ar egwyddorion, er mwyn datblygu ochr yn ochr â newid technolegol cyflym. Byddwn yn defnyddio ein gweithgareddau ar asesu a monitro a gwerthuso risgiau i barhau i asesu mesurau i atal a lliniaru unrhyw risgiau newydd a risgiau sy'n cyflymu mewn cysylltiad â deallusrwydd artiffisial, gan gynnwys y rhai a achosir o bosibl drwy ddatblygu'r systemau mwyaf pwerus, mewn ffordd gymesur, effeithiol ac wedi'i thargedu.

158. Rydym yn cytuno y bydd angen i fframwaith effeithiol ar gyfer rheoleiddio deallusrwydd artiffisial gydbwyso cysondeb ag anghenion sectorau penodol yn ofalus er mwyn cefnogi arloesi'n gyfrifol. Mae ein fframwaith sy'n canolbwyntio ar gyd-destun yn adeiladu ar arbenigedd rheoleiddwyr y DU yn eu priod feysydd, gan sicrhau bod

---

<sup>117</sup> [Frontier AI: capabilities and risks](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

diwydiannau gwahanol yn cael budd o'r wybodaeth reoleiddio sy'n bodoli eisoes. Er bod y dull gweithredu hwn yn symleiddio cydymffurfiaeth mewn sectorau penodol, rydym yn cydnabod bod angen cysondeb a chydgyssylltu rhwng rheoleiddwyr er mwyn creu tirwedd reoleiddiol hawdd ei deall i fusnesau a defnyddwyr. Dyna pam, fel y nodwn yn fanwl yn ein hymatebion i gwestiynau ynglŷn â galluoedd rheoleiddwyr a phyllau tywod a meinciau arbrofi deallusrwydd artiffisial (adrannau 6.5 a 6.10), rydym wedi bod yn canolbwyntio ar adeiladu ar gryfderau presennol rheoleiddwyr y DU drwy sefydlu gwasanaeth cyngori peilot i arloeswyr ym maes deallusrwydd artiffisial drwy'r Fforwm Cydweithredu Rheoleiddio Digidol, sy'n rhannu arweiniad ar weithredu, ac yn meithrin gallu cyffredin ymhlith rheoleiddwyr.

159. Ochr yn ochr â'n gwaith i gyflawni'r gweithgareddau asesu risg a galluoedd rheoleiddio a chydgyssylltu yn gyflym, mae'r DU wedi arwain y ffordd o ran cynnull arweinwyr y byd yn yr Uwchgynhadledd gyntaf erioed ar Ddiogelwch Deallusrwydd Artiffisial er mwyn pennu dull cyson o ymdrin â'r risgiau mwyaf dybryd sy'n gysylltiedig â thechnoleg deallusrwydd artiffisial arloesol. Cytunodd gwledydd ar Ddatganiad Bletchley yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial, gan gydnabod yr angen am gydweithio rhyngwladol i ddeall risgiau a chyfluoedd deallusrwydd artiffisial ffin eithaf.<sup>118</sup> Byddwn yn cyflwyno Adroddiad Rhyngwladol ar Wyddoniaeth Diogelwch Deallusrwydd Artiffisial a fydd yn torri tir newydd er mwyn meithrin dealltwriaeth o ddeallusrwydd artiffisial uwch sy'n seiliedig ar dystiolaeth.<sup>119</sup> Hefyd, bydd y DU, drwy'r Sefydliad Diogelwch Deallusrwydd Artiffisial, yn cydweithio â gwledydd eraill, gan gynnwys yr UD, i wella ein gallu i ymchwilio i risgiau deallusrwydd artiffisial a'u gwerthuso, gan danlinellu ein gallu i sbarduno newid drwy gydgyssylltu rhyngwladol ar y pwnc hollbwysig hwn.

160. Caiff ein gwaith yn yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial ei ategu drwy ymgysylltu amlochrog mewn fforymau eraill sy'n canolbwyntio ar ddeallusrwydd artiffisial, megis proses Hiroshima'r G7, fforwm y G20, y CU, y Bartneriaeth Fyd-eang ar Ddeallusrwydd Artiffisial, a Chyngor Ewrop. Wrth ymgymryd ag ymgysylltu amlochrog, rydym yn mynd ati i ddefnyddio cryfderau, arbenigedd, ac aelodaeth pob fforwm er mwyn atal gorgyffwrdd neu anghysondebau â systemau rheoleiddio eraill, gan sicrhau eu bod yn ychwanegu cymaint o werth â phosibl at drafodaethau byd-eang ar lywodraethu deallusrwydd artiffisial a gwerthoedd a blaenoriaethau economaidd y DU. Mae'r DU hefyd yn ceisio cydweithrediad dwyochrog â llawer o bartneriaid, sy'n adlewyrchu ein hymrwymiad i'r gallu i ryngweithredu a phennu normau rhyngwladol ar gyfer arloesi'n gyfrifol ym maes deallusrwydd artiffisial.

## 6.8. Cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial

**L1. Pa heriau a allai godi pan fydd rheoleiddwyr yn cymhwyso'r egwyddorion mewn gwahanol gymwysiadau a systemau deallusrwydd artiffisial? Sut y gallem ymateb i'r heriau hyn drwy ein fframwaith rheoleiddio deallusrwydd artiffisial arfaethedig?**

**L2.i. A ydych chi'n cytuno, drwy roi ein hegwyddorion ar waith drwy fframweithiau cyfreithiol sy'n bodoli eisoes, y byddwn yn neilltuo cyfrifoldeb am ddeallusrwydd artiffisial drwy gydol cylch oes mewn ffordd deg ac effeithiol?**

<sup>118</sup> [The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg; Y Swyddfa Dramor, y Gymanwlad a Datblygu; Swyddfa Prif Weinidog y DU, 10 Downing Street, 2023.

<sup>119</sup> [International expertise to drive International AI Safety Report](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2024.

**L.2.ii. Sut y gellid ei wella, os o gwbl?**

**L3. Os ydych yn gweithio i fusnes sy'n datblygu, yn defnyddio, neu'n gwerthu deallusrwydd artiffisial, sut rydych yn rheoli risg deallusrwydd artiffisial ar hyn o bryd, gan gynnwys drwy'r gadwyn gyflenwi ehangach? Sut y gallai'r llywodraeth gefnogi ffordd effeithiol o reoli risgiau sy'n gysylltiedig â deallusrwydd artiffisial?**

**Crynodeb o gwestiynau L1-L3:**

161. Er bod ymatebwyr yn canmol buddiannau dull gweithredu sy'n seiliedig ar egwyddorion, roedd bron eu hanner yn pryderu ynglŷn â phroblemau cydgysylltu posibl rhwng rheoleiddwyr a chysondeb rhwng sectorau. Roedd rhai yn pryderu ynglŷn â rhyngddibyniaethau dryslyd rhwng y fframwaith rheoleiddio deallusrwydd artiffisial a deddfwriaeth gyfredol. Gofynnodd ymatebwyr am ganllawiau gan reoleiddwyr sy'n seiliedig ar sectorau, adnoddau cydymffurfio, ac am i reoleiddwyr ymgysylltu â diwydiant. Nododd rhai ymatebwyr hefyd bwysigrwydd cysoni a chydweithio rhyngwladol.

162. Roedd y mwyafrif o'r ymatebwyr yn anghytuno y byddai cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial gydol cylch oes yn cael ei neilltuo mewn ffordd deg ac effeithiol drwy roi'r egwyddorion ar waith drwy fframweithiau cyfreithiol sy'n bodoli eisoes. Roedd ychydig o dan draean o'r ymatebwyr yn teimlo y dylai'r llywodraeth egluro atebolrwydd sy'n gysylltiedig â deallusrwydd artiffisial. Fodd bynnag, ni chafwyd cytundeb amlwg ynglŷn â phwy ddylai fod yn atebol, gyda'r ymatebwyr yn nodi amrywiaeth o gyfrifoldebau posibl am wahanol gyfranogwyr yn ystod cylch oes deallusrwydd artiffisial. Cydnabuwyd cymhlethdod cadwyni gwerth deallusrwydd artiffisial a'r amrywiadau posibl mewn achosion defnydd dro ar ôl tro. Roedd rhai yn mynegi pryderon ynglŷn â bylchau mewn deddfwriaeth gyfredol, gan gynnwys eiddo deallusol, gwasanaethau cyfreithiol, a chyfraith cyflogaeth.

163. Nododd tua chwarter yr ymatebwyr i L2.ii y byddai angen deddfwriaeth a phwerau rheoleiddio newydd er mwyn neilltuo atebolrwydd yn effeithiol drwy gydol y cylch oes. Pwysleisiodd ymatebwyr bwysigrwydd cael person sy'n gyfreithiol gyfrifol am ddeallusrwydd artiffisial mewn sefydliadau, gydag ychydig yn awgrymu trefniadau ar gyfer deallusrwydd artiffisial a oedd yn cyfateb i Swyddogion Diogelu Data. Roedd rhai ymatebwyr am weld mwy o fanylion ynglŷn â sut y caiff yr egwyddorion eu rhoi ar waith drwy ddeddfwriaeth gyfredol, gydag ychydig yn argymhell y byddai canllawiau rheoleiddio yn rhoi eglurder. Nododd nifer bach o ymatebwyr y byddai'r swyddogaethau canolog arfaethedig, gan gynnwys asesu risgiau, sganio'r gorwel, a monitro a gwerthuso, yn helpu i asesu ac addasu'r fframwaith er mwyn sicrhau bod y cyfrifoldeb cyfreithiol am risgiau newydd sy'n gysylltiedig â deallusrwydd artiffisial yn cael ei ddogsbarnu'n ddigonol. Hefyd, awgrymodd un neu ddau o'r ymatebwyr fesurau cyn cyflwyno megis trwyddedu a chymeradwyo cyn cyflwyno ar y farchnad.

164. Dywedodd bron hanner y sefydliadau a ymatebodd i L3 wrthym eu bod yn defnyddio prosesau asesu risg ar gyfer deallusrwydd artiffisial, gyda llawer yn adeiladu ar arferion gorau'r sector neu arweiniad gan gyrff masnach. Cyfeiriodd ymatebwyr at y fframweithiau cyfreithiol presennol sy'n cofnodi risgiau sy'n gysylltiedig â deallusrwydd artiffisial, megis cyfreithiau diogelwch cynhyrchion a diogelu data, gan bwysleisio y dylai unrhyw fesurau deallusrwydd artiffisial yn y dyfodol osgoi dyblygu neu wrthddweud rheolau sy'n bodoli eisoes. Awgrymodd ymatebwyr y byddai'n ddefnyddiol i fusnesau ddeall barn y llywodraeth ar arferion gorau deallusrwydd artiffisial, gyda rhai yn argymhell canllaw canolog ar ei ddefnyddio'n ddiogel. Gofynnodd rhai busnesau llai o faint am gymorth wedi'i dargedu i roi'r egwyddorion deallusrwydd artiffisial ar waith.

165. Pwysleisiodd ymatebwyr yn gyson pa mor bwysig yw tryloywder fel adnodd i addysgu, codi ymwybyddiaeth, cydsynio a herio. Soniodd llawer o'r ymatebwyr, gan adlewyrchu atebion i Gwestiwn 2 a F1, y dylai sefydliadau fod yn dryloyw ynglŷn â'r defnydd a wneir o ddeallusrwydd artiffisial, allbynnau, a data hyfforddi.

## Ymateb:

166. Mae'n bleser gennym nodi cefnogaeth gyffredinol ymatebwyr i ddull o reoleiddio deallusrwydd artiffisial sy'n seiliedig ar egwyddorion, sy'n gallu rhoi trosolwg cymesur o'r ffyrdd posibl niferus o gymhwyso a defnyddio technolegau deallusrwydd artiffisial. Rydym yn cytuno â'r ymatebwyr, wrth i ni roi'r fframwaith a nodwyd yn y papur gwyn ar waith, ei bod yn bwysig cydgysylltu rhwng rheoleiddwyr, sectorau, y fframweithiau cyfreithiol presennol, a'r dirwedd reoleiddio ryngwladol sy'n newid yn gyflym. Dyma pam rydym wedi mynd ati ar gyflymder i bennu gweithgareddau'r swyddogaeth ganolog a amlinellwyd yn y papur gwyn (gweler adran 5.1.2 am drosolwg manwl).

167. Nodwn fod cwestiynau o hyd ynglŷn â sut i neilltuo cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial gydol cylch oes mewn ffordd deg ac effeithiol. Rydym hefyd yn cydnabod bod llawer o'r ymatebion wedi cefnogi ymyriadau pellach gan y llywodraeth i sicrhau bod atebolrwydd am gadwyn werth deallusrwydd artiffisial yn cael ei neilltuo mewn ffordd deg ac effeithiol. Pwysleisiodd yr ymatebion pa mor gymhleth ac amrywiol yw cadwyni cyflenwi deallusrwydd artiffisial, gydag achosion defnydd yn codi cwestiynau moesegol a thechnegol eang. Rydym yn cytuno nad oes ateb hawdd o ran neilltuo cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial ac rydym hefyd yn cytuno ei bod yn bwysig cael atebolrwydd am ddeallusrwydd artiffisial yn iawn er mwyn cefnogi arloesedd ac ennyn ymddiriedaeth y cyhoedd. Gan adeiladu ar yr ymrwymiad i ystyried modelau sylfaen yn y papur gwyn, mae ein gwaith cychwynol ar atebolrwydd cylch oes wedi canolbwyntio ar systemau diben cyffredinol tra galluog (gweler adran 5.2 am fanylion).

168. Rydym hefyd yn parhau i ddadansoddi sut mae fframweithiau cyfreithiol presennol yn neilltuo atebolrwydd a chyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial gydol y cylch oes. Mae ein dadansoddiad cychwynol yn awgrymu na fydd dull o reoleiddio deallusrwydd artiffisial sy'n seiliedig ar gyd-destun yn ddigon o bosib i ymdrin â'r risgiau sy'n codi o ganlyniad i systemau diben cyffredinol tra galluog gan nad yw'r dull hwn yn neilltuo atebolrwydd i ddatblygwyr y systemau hynny. Rydym yn ystyried nifer o rwymedigaethau posibl wedi'u targedu at ddatblygwyr y systemau hyn, gan gynnwys y rhai a awgrymwyd gan ymatebwyr megis caniatadau cyn cyflwyno ar y farchnad, trwyddedu modelau, fframweithiau atebolrwydd a llywodraethu, mesurau tryloywder, a newidiadau i fframweithiau cyfreithiol presennol. Wrth i ni barhau i iteru'r fframwaith rheoleiddio deallusrwydd artiffisial, byddwn yn ystyried cyflwyno mesurau i neilltuo atebolrwydd yn effeithiol a dosbarthu cyfrifoldeb cyfreithiol yn deg i'r rhai yn y cylch oes sydd yn y sefyllfa orau i liniaru risgiau sy'n gysylltiedig â deallusrwydd artiffisial.

169. Mae'n galonogol i ni nodi'r amrywiaeth eang o brosesau asesu a rheoli risg y dywedodd ymatebwyr eu bod eisoes yn eu defnyddio. Mae ein papur ar "Emerging processes for frontier AI safety" yn amlinellu set o arferion i lywio'r broses o ddatblygu polisiâu diogelwch deallusrwydd artiffisial sefydliadol.<sup>120</sup> Mae'n cynnig ciplun o syniadau addawol ac arferion cysylltiedig ym maes diogelwch deallusrwydd artiffisial heddiw. Fel y'i trafodwyd mewn ymateb i gwestiynau ynglŷn â'r egwyddorion traws-sector (adran 6.1), rydym yn cydnabod y gefnogaeth gyffredinol i fesurau ar dryloywder a byddwn yn parhau

---

<sup>120</sup> [Emerging Processes for Frontier AI Safety](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023

â'n gwaith yn asesu a yw mesurau yn rhoi tryloywder i ddefnyddwyr terfynol deallusrwydd artiffisial a'r rhai sy'n rhan o gylch oes deallusrwydd artiffisial a pha rai sy'n cynnig y tryloywder mwyaf ystyrlon.

## 6.9. Modelau sylfaen a'r fframwaith rheoleiddio

**F1. Pa heriau penodol y bydd modelau sylfaen megis modelau iaith mawr neu fodelau ffynhonnell agored yn eu hachosi i reoleiddwyr sy'n ceisio pennu cyfrifoldeb cyfreithiol am ganlyniadau deallusrwydd artiffisial?**

**F2. A ydych chi'n cytuno bod mesur lefelau cyfrifiannu yn cynnig adnodd posibl y gellid ei ystyried fel rhan o drefniadau llywodraethu modelau sylfaen?**

**F3. A oes unrhyw ddulliau eraill o lywodraethu modelau sylfaen a fyddai'n fwy effeithiol?**

Crynodeb o gwestiynau F1-F3:

170. Er bod ymatebwyr yn cefnogi'r fframwaith rheoleiddio deallusrwydd artiffisial a nodwyd yn y papur gwyn, roedd llawer yn pryderu y gallai fod angen dull rheoleiddio pwrpasol ar gyfer modelau sylfaen. Nododd rhai ymatebwyr mai prif nodwedd modelau sylfaen yw eu cymhlethdod technegol, gan bwysleisio eu potensial i ategu llawer o gymwysiadau gwahanol mewn nifer o sectorau. Pwysleisiodd bron chwarter yr ymatebwyr fod modelau sylfaen yn ei gwneud yn anodd i bennu cyfrifoldeb cyfreithiol am ganlyniadau deallusrwydd artiffisial, gan rannu achosion defnydd damcaniaethol lle roedd bai ar y rhai i fyny'r gadwyn ac i lawr y gadwyn. Pwysleisiodd ymatebwyr fod diffyg tryloywder technegol, cadwyni cyflenwi cymhleth, ac anghymesureddau o ran gwybodaeth yn atal esboniadwyedd digonol, atebolrwydd, ac asesiadau risg ar gyfer modelau sylfaen.

171. Mynegodd tua un o bob pum ymatebydd bryderon ynglŷn â sut mae modelau sylfaen yn defnyddio data, gan gynnwys a yw'r data o ansawdd digon da, yn briodol ar gyfer cymwysiadau i lawr y gadwyn, yn cydymffurfio â'r gyfraith bresennol, ac yn dod o ffynonellau moesegol. Nododd rhai nad yw'n glir pwy sy'n gyfrifol am benderfynu a yw data yn briodol ar gyfer cymhwysiad penodol. Pwysleisiodd ymatebwyr nad oes diffiniad clir o ddata hyfforddi, safonau technegol, na mesuriadau meincnodi.

172. Nododd rhai ymatebwyr bryderon ynglŷn â mynediad ehangach at ddeallusrwydd artiffisial, gan gynnwys ffynhonnell agored, rhyddhau gwybodaeth yn amhriodol, neu ddefnydd maleisus o fodelau. Fodd bynnag, nododd nifer tebyg o ymatebwyr bwysigrwydd ffynhonnell agored i arloesi, tryloywder, ac ymddiriedaeth ym maes deallusrwydd artiffisial.

173. Roedd hanner yr ymatebwyr yn teimlo bod cyfrifiannu yn fesur dirprwyol annigonol ar gyfer gofynion llywodraethu, gyda rhai yn argymhell asesu modelau yn ôl eu galluoedd a'u cymwysiadau yn lle hynny. Roedd ymatebwyr yn teimlo y byddai mesurau dilysu modelau, megis archwiliadau a gwerthusiadau, yn effeithiol, gyda rhai yn awgrymu y dylai'r rhain fod yn ofynion gorfodol. Nododd ychydig bwysigrwydd monitro i lawr y gadwyn neu drefniadau gwylidwriaeth ar ôl mynd ar y farchnad.

174. Roedd tua thraean o'r ymatebwyr yn cefnogi mesurau llywodraethu, gan gynnwys adnoddau ar gyfer deallusrwydd artiffisial dibynadwy megis safonau technegol a sicrwydd. Awgrymodd un ymatebydd bwl tywod cyn cyflwyno. Roedd ychydig o'r ymatebwyr yn

cefnogi moratoria, gwaharddiadau, neu gyfyngiadau. Awgrymodd nifer bach o ymatebwyr y gellid defnyddio contractau, trwyddedau, cytundebau â defnyddwyr a mesurau (seiber) diogelwch i lywodraethu modelau sylfaen.

## Ymateb:

175. Rydym yn cydnabod yr amrywiol heriau y mae'r ymatebwyr wedi eu codi o ran modelau sylfaen ac yn nodi'r sylw penodol ar nodweddion craidd modelau sylfaen megis diffyg tryloywder technegol a chymhlethdod. Rydym hefyd yn cydnabod bod heriau yn codi am y gall modelau sylfaen fod yn eang eu cymwysiadau posibl ac, fel y cyfryw, dorri ar draws sectorau ac effeithio ar amrywiaeth o risgiau. Mae ein dadansoddiad yn dangos y gall llawer o reoleiddwyr wynebu anhawster i orfodi'r rheolau a'r cyfreithiau presennol ar ddatblygwyr systemau deallusrwydd artiffisial diben cyffredinol tra galluog o fewn eu cylchoedd gwaith statudol presennol mewn ffordd sy'n lliniaru risg yn effeithiol.

176. Mewn ymateb i geisiadau mynych am ymyriadau rheoleiddio penodol wedi'u targedu at fodolau sylfaen, rydym wedi bod yn ystyried effaith modelau sylfaen ar atebolrwydd cylch oes am ddeallusrwydd artiffisial. Yn y papur gwyn ar reoleiddio deallusrwydd artiffisial, nodwyd gennym y dylai cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial fod ar y rhai sydd yn y sefyllfa orau i liniaru unrhyw risgiau posibl sy'n codi. Mae ein hasesiad yn awgrymu, er gwaethaf eu gallu i liniaru risgiau wrth ddylunio a datblygu deallusrwydd artiffisial, fod y rheolau a'r cyfreithiau presennol yn annhebygol o effeithio ar y sefydliadau sy'n adeiladu systemau diben cyffredinol tra galluog mewn ffordd sy'n ddigonol i liniaru risg. Dyna pam rydym yn ystyried opsiynau ar gyfer ymyriadau cymesur, wedi'u targedu sy'n canolbwyntio ar y systemau hyn a'r risgiau sy'n codi o ganlyniad iddynt. Rydym wedi bod yn asesu mesurau i liniaru risg yn ystod prosesau dylunio, hyfforddi, a datblygu systemau diben cyffredinol tra galluog. Rydym hefyd wedi bod yn ystyried opsiynau i sicrhau atebolrwydd effeithiol, gan gynnwys rhwymedigaethau cyfreithiol, ond osgoi biwrocratiaeth feichus ar yr un pryd.

177. Nodwn farn ymatebwyr bod cyfrifiannu yn fesur dirprwyol amherffaith am allu modelau sylfaen. Fel rhan o'n gwaith yn ystyried y canllawiau cywir ar gyfer systemau diben cyffredinol tra galluog, rydym yn archwilio'r ffordd orau o gwmpasu unrhyw ofynion rheoleiddio yn seiliedig ar alluoedd modelau, a'r risgiau sy'n gysylltiedig â'r rhain, lle bynnag y bo modd. Ond rydym yn cydnabod, mewn rhai achosion, efallai y bydd angen rhoi rheolaethau ar waith cyn bod gallu model yn hysbys. Yn yr achosion hyn, efallai y bydd angen defnydd cyfyngedig a gofalus o fesurau dirprwyol er mwyn targedu gofynion rheoleiddio dim ond at y systemau hynny sy'n achosi'r risgiau posibl mwyaf sylweddol. Mae ein dadansoddiad rhagarweiniol yn dangos y gallai trothwyon cychwynnol fod yn seiliedig ar ragolygon o alluoedd drwy ddefnyddio cyfuniad o ddau fesur dirprwyol: cyfrifiannu a meincnodi galluoedd. Fodd bynnag, efallai y bydd angen ystod o drothwyon. Gweler adran 5.2 am ragor o fanylion.

178. Er mwyn rhoi mwy o eglurder ynglŷn ag arferion gorau ar gyfer arloesi'n gyfrifol ym maes deallusrwydd artiffisial – gan gynnwys defnyddio data – gwnaethom gyhoeddi set o brosesau diogelwch sy'n cael eu datblygu ar gyfer cwmnïau deallusrwydd artiffisial ffin eithaf ar gyfer yr Uwchgynhadledd ar Ddiogelwch Deallusrwydd Artiffisial yn 2023.<sup>121</sup> Mae'r ddogfen yn cyfuno syniadau datblygol ynglŷn â diogelwch deallusrwydd artiffisial ac mae wedi cael ei hysgrifennu i sefydliadau deallusrwydd artiffisial a'r rhai sydd am ddeall eu polisiâu diogelwch yn well. Byddwn yn diweddarau'r canllaw hwn erbyn diwedd y flwyddyn ac rydym yn parhau i annog cwmnïau deallusrwydd artiffisial i ddatblygu arferion gorau (gweler adran 5.2.2 am fanylion).

---

<sup>121</sup> [Emerging Processes for Frontier AI Safety](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023

179. Rydym yn cydnabod barn ymatebwyr ar werth a risgiau deallusrwydd artiffisial ffynhonnell agored. Gall mynediad agored gynnig manteision eang, gan gynnwys helpu i liniaru rhai o'r risgiau a achosir gan systemau diben cyffredinol tra galluog. Fodd bynnag, gall trefniadau rhyddhau agored hefyd gynyddu'r risg o gamddefnydd. Credwn y dylai pob system bwerus, sy'n beryglus o bosibl, gael ei hasesu o ran risgiau cyn cael ei rhyddhau. Byddwn yn parhau i fonitro ac asesu effeithiau mynediad model agored ar risg. Byddwn hefyd yn ystyried yn ofalus effaith unrhyw fesurau posibl i reoleiddio systemau ffynhonnell agored ar gystadleuaeth, arloesi, a lliniaru risgiau yn fwy cyffredin.

180. Fel y'i nodwyd yn adran 5.2, byddwn yn parhau â'n dadansoddiad o bolisi technegol er mwyn mireinio ein syniadau ynglŷn â systemau diben cyffredinol tra galluog yng nghydestun rheoleiddio deallusrwydd artiffisial ac atebolrwydd cylch oes. Byddwn yn parhau i ymgysylltu ag arbenigwyr allanol ar nifer o bynciau heriol megis pa mor effeithiol y gallai mesurau gwirfoddol fod o ran lliniaru risgiau a chwmpas cywir unrhyw ymyriadau rheoleiddiol ychwanegol, gan gynnwys mesurau dirprwyol a throthwyon galluoedd. Byddwn hefyd yn parhau i ystyried cwestiynau ynglŷn ag atebolrwydd, gan gynnwys i ba raddau y gall cyfreithiau a rheoleiddwyr presennol "ymestyn" drwy'r gadwyn werth i dargedu datblygwyr systemau diben cyffredinol tra galluog ac effaith bosibl trefniadau rhyddhau agored. Byddwn hefyd yn ymgysylltu â rheoleiddwyr er mwyn dysgu oddi wrth y gwaith y maent eisoes wedi'i wneud ar y pwnc hwn. Er enghraifft, byddwn yn parhau i ymgysylltu â'r Awdurdod Cystadleuaeth a Marchnadoedd ar ei waith ar fodolau sylfaen.

## 6.10. Pyllau tywod a meiciau arbrofi deallusrwydd artiffisial

**S1. I ba raddau y byddai'r modelau pwll tywod a ddisgrifir yn adran 3.3.4 yn cefnogi arloesedd?**

**S2. Beth y gallai'r llywodraeth ei wneud i sicrhau manteision mwyaf posibl pyllau tywod i arloeswyr deallusrwydd artiffisial?**

**S3. Beth y gallai'r llywodraeth ei wneud i hwyluso cyfranogiad mewn pwll tywod rheoleiddio deallusrwydd artiffisial?**

**S4. Pa sectorau diwydiannol neu ddosbarthiadau o gynnyrch a fyddai'n cael y budd mwyaf o bwll tywod deallusrwydd artiffisial?**

Crynodeb o gwestiynau S1-S4:

181. Ar y cyfan, roedd yr ymatebwyr yn gefnogol iawn i bwll tywod rheoleiddio ar gyfer deallusrwydd artiffisial. Roedd y gyfran fwyaf o ymatebwyr yn cytuno mai modelau pwll tywod "amlssector, amlreoleiddiwr" a "sector sengl, amlreoleiddiwr" fyddai fwyaf tebygol o gefnogi arloesedd, gan nodi y byddai'r sail draws-sector neu draws-reoleiddiol yn helpu i ddatblygu canllawiau effeithiol mewn ymateb i faterion byw, cysoni rheolau, a chydgyssylltu'r gwaith o roi'r fframwaith rheoleiddio deallusrwydd artiffisial ar waith. Er nad oedd consensws mwyafrifol ar sector penodol a fyddai'n cael y budd mwyaf o bwll tywod, nododd y gyfran fwyaf o ymatebwyr i'r cwestiwn mai gofal iechyd a dyfeisiau meddygol fyddai'n cael y budd mwyaf o bwll tywod deallusrwydd artiffisial, ac yna wasanaethau ariannol a thrafnidiaeth.

182. Awgrymodd rhai ymatebwyr y dylid cydweithio ag ecosystem ehangach deallusrwydd artiffisial er mwyn sicrhau budd mwyaf posibl pyllau tywod i arloeswyr deallusrwydd artiffisial. Argymhellodd llawer y dylid adeiladu ar gryfderau presennol tirwedd reoleiddio'r



DU, megis y Fforwm Cydweithredu Rheoleiddio Digidol. Yn gysylltiedig â hyn, nododd ychydig o'r ymatebwyr fod pwll tywod rheoleiddio deallusrwydd artiffisial yn cynnig cyfle i'r DU ddangos arweinyddiaeth fyd-eang ym maes rheoleiddio deallusrwydd artiffisial a safonau technegol drwy rannu canfyddiadau ac arferion gorau yn rhyngwladol.

183. Argymhellodd rhai ymatebwyr y dylid gwneud gwybodaeth yn hygyrch er mwyn sicrhau manteision mwyaf y pwll tywod i gyfranogwyr ac ecosystem ehangach deallusrwydd artiffisial. Roedd ymatebwyr am i lwybrau cyfranogi, hyfforddiant, ac adnoddau eraill fod yn hygyrch, yn dechnegol ac yn ariannol. Nododd llawer o'r ymatebwyr y byddai canllawiau ac adnoddau hygyrch yn ei gwneud yn bosibl i sefydliadau gymryd rhan yn y pwll tywod. Yn benodol, pwysleisiodd ymatebwyr fanteision gwybodaeth hygyrch i fusnesau bach ac egin fusnesau nad ydynt yn gyfarwydd â'r broses reoleiddio. Dadleuodd ymatebwyr o blaid adroddiadau rheolaidd ar brosesau pwll tywod, tystiolaeth, canfyddiadau, a chanlyniadau er mwyn annog arferion gorau "busnes arferol" ar gyfer deallusrwydd artiffisial yn yr ecosystem ehangach.

184. Nododd ymatebwyr bwysigrwydd lleihau'r baich gweinyddol ar fusnesau bach ac egin fusnesau er mwyn sicrhau llai o rwystr i fynediad i'r rhai sydd â llai o adnoddau sefydliadol. Nododd rhai y byddai cymorth ariannol yn helpu i sicrhau y gallai busnesau bach ac egin fusnesau gymryd rhan mewn ymchwil a datblygu ym maes pyllau tywod deallusrwydd artiffisial sy'n gofyn am lawer o adnoddau. Roedd ymatebwyr yn teimlo y byddai rhannu tystiolaeth, canllawiau, ac adnoddau yn sicrhau bod ecosystem ehangach deallusrwydd artiffisial yn cael budd o'r pwll tywod. Awgrymodd rhai y byddai mynediad at setiau data neu gynlluniau achredu cynnyrch yn rhoi cymhelliant i gymryd rhan mewn pyllau tywod amgylchedd prawf o dan oruchwyliaeth.

## Ymateb:

185. Mae'r ymateb i'r ymgynghoriad – sy'n gyson ag ymchwil annibynnol a gomisiynwyd drwy Gronfa Arloeswyr y Rheoleiddwyr – wedi helpu i lywio penderfyniad y Llywodraeth i ariannu gwasanaeth cynghori peilot â sawl rheoleiddiwr a gynigir gan y Fforwm Cydweithredu Rheoleiddio Digidol: yr Hwb Deallusrwydd Artiffisial a Digidol. Yn benodol, mae wedi helpu i'w gwneud yn glir mai gwasanaeth rheoleiddio newydd sy'n debygol o ychwanegu'r gwerth mwyaf o ran helpu arloeswyr deallusrwydd artiffisial o amryw sector i ddeall y cyfundrefnau rheoleiddio niferus sy'n llywodraethu'r defnydd o gynhyrchion a gwasanaethau deallusrwydd artiffisial trawsbynciol, yn hytrach na thrwy dargedu un cylch gwaith rheoleiddio neu un sector rheoleiddedig penodol.

186. Mae Hwb Deallusrwydd Artiffisial a Digidol y Fforwm Cydweithredu Rheoleiddio Digidol yn dwyn ynghyd bedwar o'r rheoleiddwyr deallusrwydd artiffisial a thechnolegau digidol mwyaf allweddol, gan gynnwys yr Awdurdod Cystadleuaeth a Marchnadoedd, Swyddfa'r Comisiynydd Gwybodaeth, Ofcom, a'r Awdurdod Ymddygiad Ariannol. Gyda'i gilydd, mae'r rheoleiddwyr hyn yn gyfrifol am oruchwyllo rhai o'r cyfundrefnau rheoleiddio mwyaf arwyddocaol sy'n llywodraethu cynhyrchion deallusrwydd artiffisial, boed hynny ar draws yr economi (diogelu data, rheoleiddio cystadleuaeth a defnyddwyr) neu mewn sectorau (gwasanaethau ariannol, telathrebu a darlledu).

187. Pwysleisiodd ymatebwyr i'r ymgynghoriad hefyd pa mor bwysig ydyw bod gwybodaeth ac adnoddau sy'n ymwneud â'r pwll tywod yn hygyrch er mwyn sicrhau'r manteision mwyaf posibl. Nododd ymatebwyr fod angen lleihau'r baich cydymffurfio ar fusnesau bach ac egin fusnesau yn benodol. Unwaith eto, mae'r ystyriaethau hyn yn ganolog i'r gwaith o ddylunio a gweithredu Hwb Deallusrwydd Artiffisial a Digidol y Fforwm Cydweithredu Rheoleiddio Digidol. Yn ogystal â chynnig cymorth wedi'i deilwra i arloeswyr sy'n cymryd rhan, a fydd ar gael drwy broses ymgeisio ar-lein syml, bydd yr Hwb hefyd yn cyhoeddi astudiaethau achos wedi'u hanonymio a chanllawiau i gefnogi cronfa ehangach o arloeswyr sy'n

wynebu heriau tebyg o ran cydymffurfiaeth. Mae ein hymchwil wedi awgrymu y bydd storfa o achosion defnydd fel hon yn fodd effeithiol iawn i ehangu cyrhaeddiad ac effaith cynllun peilot o'r fath.

188. Nodwn fod rhai ymatebwyr wedi awgrymu y byddai cymhellion ychwanegol megis achredu cynhyrchion neu fynediad at ddata yn annog busnesau i gymryd rhan mewn pwll tywod ar gyfer deallusrwydd artiffisial. Byddai'r cymhellion ychwanegol hyn yn fwyaf addas ar gyfer model pwll tywod amgylchedd prawf o dan oruchwyliaeth. Gan y bydd cam peilot Hwb Deallusrwydd Artiffisial a Digidol y Fforwm Cydweithredu Rheoleiddio Digidol yn canolbwyntio ar roi cymorth i gydymffurfio, ni fydd y cymhellion ychwanegol hyn yn cael eu cynnwys. Fodd bynnag, rydym yn ymrwymedig i adolygu sut mae angen i'r gwasanaeth ddatblygu – a pha fesurau eraill sydd eu hangen i gefnogi arloeswyr deallusrwydd artiffisial a digidol – yng ngoleuni canfyddiadau'r cynllun peilot ac adborth pellach gan randdeiliaid.

# Atodiad A: Dull ac ymgysylltu

## Dull ymgynghori a chrynodeb o'r ymgysylltu

1. Yn sgil cyhoeddi'r papur gwyn ar reoleiddio deallusrwydd artiffisial ar 29 Mawrth 2023, gwnaethom gynnal ymgynghoriad cyhoeddus 12 wythnos ffurfiol a ddaeth i ben ar 21 Mehefin 2023. Clywsom gan fwy na 545 o unigolion a sefydliadau gwahanol i gyd.
2. Gofynnwyd i randdeiliaid gyflwyno tystiolaeth mewn ymateb i 33 o gwestiynau ar gynigion polisi'r llywodraeth ar gyfer fframwaith rheoleiddio deallusrwydd artiffisial. Gofynnwyd i randdeiliaid gyflwyno tystiolaeth drwy arolwg ar-lein, dros e-bost, neu drwy'r post. Cawsom 409 o ymatebion ysgrifenedig i gyd. Ar ôl cael gwared ar 50 o ymatebion dyblyg a gwag cafwyd 359 o ymatebion ysgrifenedig. Gweler **Ymatebion ysgrifenedig a gyflwynwyd** isod am ragor o fanylion.
3. Gwnaethom hefyd fynd ati i ymgysylltu â 364 o unigolion drwy ddigwyddiadau bord gron, gweithdai technegol, cyfarfodydd dwyochrog, a rhaglen o ymgysylltu parhaus â rheoleiddwyr. Yn ein digwyddiadau bord gron aethpwyd ati i geisio barn rhanddeiliaid a glywir ganddynt yn llai aml ac roedd y pynciau yn cynnwys effaith deallusrwydd artiffisial ar gymunedau wedi'u hymyleiddio, ymddiriedaeth ymhlith y cyhoedd, a safbwyntiau dinasyddion. Gwnaethom hefyd gynnal digwyddiadau bord gron a oedd yn canolbwyntio ar fusnesau bach a'r gymuned ffynhonnell agored. Ceir rhagor o fanylion yn yr adrannau **Dull ymgysylltu** a **Canfyddiadau ymgysylltu** isod.

## Y dull a ddefnyddiwyd i ddadansoddi ymatebion ysgrifenedig a gyflwynwyd

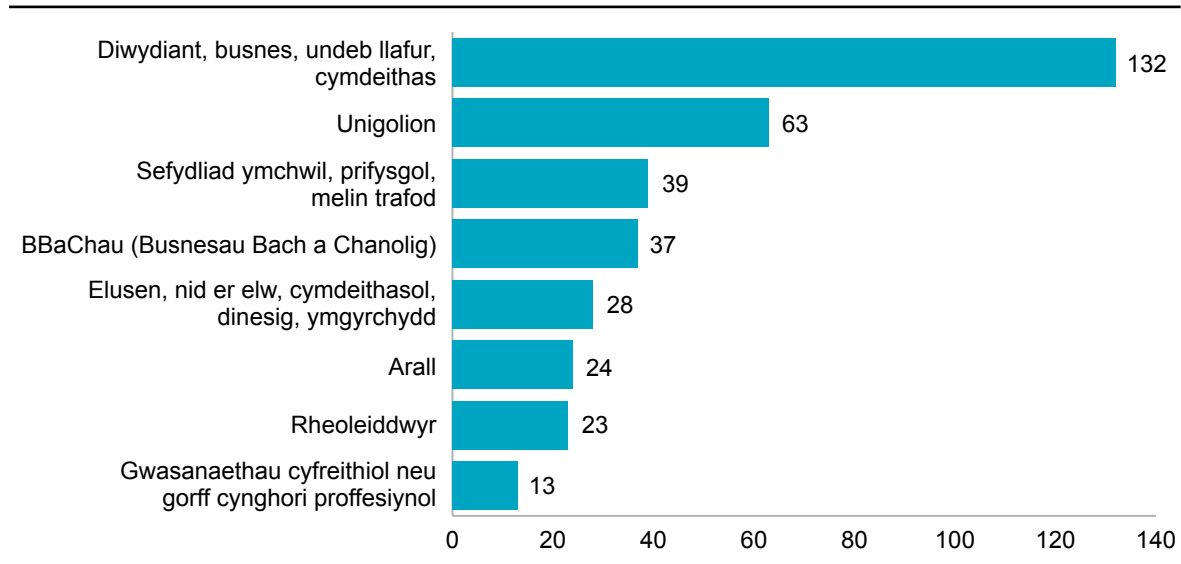
4. Cawsom ymatebion ysgrifenedig i'r ymgynghoriad gan sefydliadau ac unigolion drwy arolwg ar-lein ac e-bost. O'r cyfanswm o 409 o ymatebion, cawsom 232 drwy ein harolwg ar-lein a 177 drwy e-bost.
5. O blith y 33 o gwestiynau, roedd 12 yn gwestiynau caeedig a oedd yn cynnig opsiynau ymateb rhagnodedig ar yr arolwg ar-lein. Aethom ati i godio â llaw ymatebion a gyflwynwyd drwy e-bost a oedd yn ymateb yn benodol i'r cwestiynau caeedig hyn er mwyn dilyn strwythur graddfa Likert. Roedd y 21 o gwestiynau eraill yn gofyn am ymatebion ansoddol testun rhydd a chafodd pob ymateb ei ddadansoddi'n unigol a'i godio â llaw. Fel y cyfryw, mae'r dadansoddiad meintiol yn cynrychioli pob rhanddeiliad a atebodd gwestiwn penodol drwy e-bost neu'r arolwg ar-lein. Nid atebwyd pob cwestiwn gan bob ymatebydd ac rydym yn cyflwyno ein canfyddiadau fel cyfran fras o'r ymatebion i'r cwestiwn.
6. Yn unol â'n hysbysiad preifatrwydd<sup>122</sup> a chytundeb preifatrwydd yr arolwg ar-lein, dim ond enwau'r unigolion a'r sefydliadau hynny a gyflwynodd dystiolaeth drwy ein harolwg ar-lein ac a gydsyniodd i'n cytundeb preifatrwydd a gyhoeddwyd yn rhestr o ymatebwyr (gweler Atodiad B).
7. Gwnaeth ymatebwyr i'r arolwg ar-lein ddewis math o sefydliad a sector eu hunain. Gwnaethom neilltuo math o sefydliad a sector â llaw ar gyfer ymatebwyr a gyflwynodd dystiolaeth ysgrifenedig drwy e-bost. Ar ôl cael gwared ar ymatebion gwag a dyblyg,

---

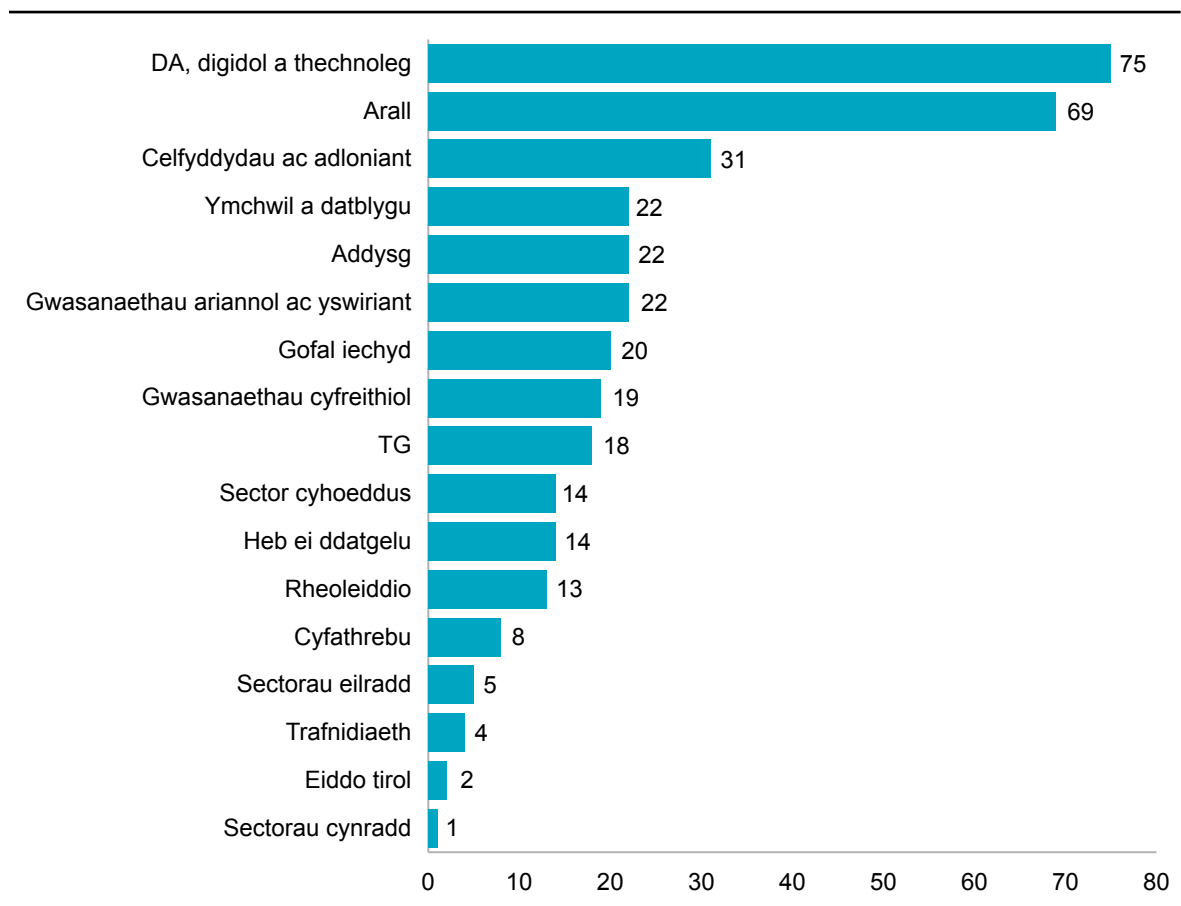
<sup>122</sup> [Office for Artificial Intelligence – information collection and analysis: privacy notice](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

cawsom ymatebion gan 8 math o sefydliad a 18 o sectorau. Mae Siart M1 yn dangos y niferoedd a ymatebodd yn ôl math o sefydliad. Daeth y mwyafrif o'r ymatebion oddi wrth ddiwydiant, busnesau, undebau llafur, a chymdeithasau masnach. Y grŵp nesaf oedd unigolion nad oeddent yn cynrychioli sefydliad ac yna grwpiau ymchwil, prifysgolion, a melinau trafod.

### Siart M1: Ymatebwyr i'r ymgynghoriad ar y papur gwyn ar reoleiddio deallusrwydd artifisial yn ôl math o sefydliad



### Siart M2: Ymatebwyr i'r ymgynghoriad ar y papur gwyn ar reoleiddio deallusrwydd artifisial yn ôl sector



*Nodyn M2: Mae sectorau cynradd yn cynnwys echdynnu deunyddiau crai, ffermio a physgota. Mae sectorau eilradd yn cynnwys cyfleustodau, adeiladu, a gweithgynhyrchu.*

8. Mae'r dadansoddiad yn ôl sector yn Siart M2 yn dangos mai gan y diwydiant deallusrwydd artiffisial, digidol a thechnoleg y cafwyd y nifer mwyaf o ymatebion. Y grŵp nesaf oedd ymatebwyr a ddewisodd "arall" ac yna'r rhai yn sector y celfyddydau ac adloniant. Mae dadansoddiad pellach o ymatebion "arall" yn awgrymu bod yr ymatebion hyn yn aml wedi cael eu cyflwyno gan unigolion nad oeddent yn cynrychioli sefydliad, gan gynnwys myfyrwyr.

9. Fel y dengys y wybodaeth ddemograffig hon, mae'n bosibl nad yw'r sampl hon, fel yn achos samplau pob ymgynghoriad ysgrifenedig, yn gynrychioliadol o farn y cyhoedd am fod rhai grwpiau wedi'u gorgynrychioli neu eu tangynrychioli.

10. Yn benodol, nodwn fod ymatebion gan nifer o randdeiliaid yn y diwydiannau creadigol yn unfath neu'n debyg iawn. Roedd yr ymatebion hyn yn canolbwyntio'n bennaf ar ddeallusrwydd artiffisial a hawlfraint. Cafodd yr ymatebion hyn eu dadansoddi a'u cynnwys yn yr un ffordd â'r holl ymatebion eraill.

11. Roedd 89 o ddarnau o dystiolaeth a anfonwyd dros e-bost yn dilyn strwythur cwestiynau ein harolwg ar-lein. Cafod y rhain eu dadansoddi ochr yn ochr ag ymatebion o'r arolwg er mwyn llywio'r dadansoddiad meintiol. Ar ôl cael gwared ar ymatebion dyblyg, gwnaethom gynnwys 66 o ymatebion a anfonwyd dros e-bost yn ein dadansoddiad.

12. Roedd 88 o'r ymatebion a anfonwyd dros e-bost yn rhoi tystiolaeth a oedd y tu hwnt i gwmpas cwestiynau ein hymgynghoriad neu nid oeddent yn cyfeirio'n benodol at y cwestiynau. Gwnaethom ddadansoddi'r ymatebion hyn a gyflwynwyd yn unigol. Er bod ein canfyddiadau o'r dadansoddiad hwn yn llywio ein hymateb cyffredinol, nid ydym yn cynnwys yr ymatebion hyn yn ein dadansoddiad meintiol am nad ydynt yn rhoi ateb penodol i gwestiynau ein hymgynghoriad. Lle y bo'n berthnasol, rydym wedi defnyddio dirnadaethau o'r ymatebion i lywio ein crynodebau ansoddol o'r cwestiynau. Ar ôl cael gwared ar ymatebion dyblyg, gwnaethom gynnwys 84 o'r rhain yn ein dadansoddiad ansoddol.

13. Cawsom 33 o ymatebion dyblyg a anfonwyd ddwywaith drwy'r arolwg ar-lein neu dros e-bost. Gofynnwyd i ni ddileu pedwar o'r ymatebion dyblyg hyn ar y sail eu bod yn anghywir ac wedi cael eu disodli gan ymateb diweddarach. Cafwyd gwared ar yr ymatebion dyblyg yn gyfan gwbl yn y dadansoddiad. Roedd y 29 o ymatebion dyblyg eraill yn rhai a anfonwyd drwy'r arolwg ar-lein a thros e-bost. Lle y bo'n briodol, cafwyd gwared naill ai ar yr e-bost neu'r ymateb i'r arolwg yn ein dadansoddiad meintiol er mwyn osgoi ystumio cyfrifon drwy ymatebion dyblyg. Fodd bynnag, er mwyn i'r manylion ychwanegol a roddwyd gael eu hystyried, gwnaethom ddadansoddi'r ddau ymateb er mwyn cyd-blethu unrhyw ddirnadaethau ychwanegol â'n dadansoddiad ansoddol cyffredinol. Diystyrwyd 17 o ymatebion ysgrifenedig eraill yn gyfan gwbl yn y dadansoddiad am eu bod yn wag neu'n cynnwys sbam. Ar ôl adolygu a chroeswrio ymatebion, diystyrwyd 50 o ymatebion ysgrifenedig a gyflwynwyd yn y dadansoddiad terfynol er mwyn osgoi gorgyfrif ymatebion gwag, dyblyg, a sbam. Roedd hynny'n gadael 359 o ymatebion, gyda 209 ohonynt wedi cael eu hanfon drwy'r arolwg ar-lein a 150 wedi cael eu hanfon dros e-bost.

14. Rydym yn defnyddio iaith ansoddol eglurhaol megis "llawer", "rhai", ac "ychydig" i grynhoi'r ymatebion ysgrifenedig a gawsom i'n hymgynghoriad. Bwriedir i'r disgrifiadau hyn awgrymu i ba raddau y codwyd thema neu deimlad penodol gan ymatebwyr. Nid atebwyd pob cwestiwn gan bob ymatebydd. Cyfeiriwn at y niferoedd bras o ymatebwyr i bob cwestiwn, gan gynnwys "hanner", "chwarter", neu "traean". Defnyddiwn y termau "bron pob un" neu "y rhan fwyaf" pan wnaeth y mwyafrif sylweddol o'r ymatebwyr gyflwyno dadl benodol neu rannu teimlad. Defnyddiwn y termau "mwyafrif" neu "dros hanner" i ddangos pan gafodd y pwynt ei rannu gan fwy na 50% o'r ymatebwyr. Defnyddiwn "llawer" pan godwyd pwynt tebyg gan lawer o'r ymatebwyr ond nad oedd dros hanner yr ymatebwyr

wedi rhannu'r thema/teimlad. Defnyddiwn "rhai" i nodi pan rannwyd thema neu deimlad gan rhwng un o bob 10 ac un o bob pum ymatebydd. Defnyddiwn "ychydig" pan wnaeth nifer llai o ymatebwyr bwynt tebyg. Defnyddiwn "nifer bach" i nodi bod pwynt wedi cael ei godi gan lai na 10 o ymatebwyr, gan nodi'n benodol os mai "un" neu "dau" ("un neu ddau") sydd dan sylw.

## Dull ymgysylltu

15. Cynhaliwyd 19 o ddigwyddiadau bord gron a oedd yn cynnwys 278 o unigolion yn cynrychioli amrywiaeth o safbwyntiau a mathau o sefydliad, gan gynnwys y diwydiant deallusrwydd artiffisial, digidol, a sefydliadau technoleg, busnesau bach ac egin fusnesau, cwmnïau sy'n defnyddio deallusrwydd artiffisial, y gymuned ffynhonnell agored, cyrff masnach ac undebau, gwasanaethau cyfreithiol, gwasanaethau ariannol, y diwydiannau creadigol, academyddion, melinau trafod, sefydliadau ymchwil, rheoleiddwyr, adrannau'r llywodraeth, y sector cyhoeddus, elusennau a grwpiau eirioli, dinasyddion, cymunedau wedi'u hymyleiddio, a chymdeithas sifil yn fwy cyffredinol.

16. Roedd rhai digwyddiadau bord gron yn canolbwyntio ar wrando gan reoleiddwyr neu randdeiliaid mewn sector penodol, gan gynnwys addysg, trafndiaeth, gwasanaethau ariannol, gwasanaethau cyfreithiol, ac iechyd a gofal cymdeithasol. Roedd eraill yn canolbwyntio ar elfennau technegol o'r fframwaith rheoleiddio megis dulliau ar gyfer dilysu deallusrwydd artiffisial, atebolrwydd, ac adnoddau ar gyfer deallusrwydd artiffisial dibynadwy, gan gynnwys safonau technegol. Roedd rhai trafodaethau wedi'u cynllunio i'n helpu i ddeall safbwyntiau rhanddeiliaid a glywn yn llai aml o bosibl: ystyriodd un effaith deallusrwydd artiffisial ar gymunedau wedi'u hymyleiddio, archwiliodd un arall rôl ymddiriedaeth y cyhoedd, canolbwyntiodd dau arall ar safbwyntiau busnesau bach a'r gymuned ffynhonnell agored, a chadeiriodd y Gweinidog dros Ddeallusrwydd Artiffisial ac Eiddo Deallusol, yr Is-iarll Camrose, ford gron i ddinasyddion yn ystod Wythnos Dechnoleg Llundain. Roedd pynciau eraill yn cynnwys diogelwch deallusrwydd artiffisial, y gallu i ryngweithredu'n rhyngwladol, dulliau o arloesi'n gyfrifol ym maes deallusrwydd artiffisial mewn diwydiant, a Chenhadaeth Technoleg Deallusrwydd Artiffisial UKRI.

17. Rydym yn ddiolchgar i'r partneriaid a weithiodd gyda ni i drefnu digwyddiadau bord gron a gweithdai, gan gynnwys y Ganolfan Moeseg Data ac Arloesi, yr Adran Addysg, yr Adran Iechyd a Gofal Cymdeithasol, yr Adran Drafndiaeth, y Weinyddiaeth Gyfiawnder, Ymchwil ac Arloesi yn y DU (UKRI), Cymdeithas Gyfrifiadurol Prydain, Hogan Lovells, Innovate Finance, Sefydliad Ada Lovelace, Sefydliad Alan Turing, Open UK, y Sefydliad Safonau Prydeinig, ac ART-AI Prifysgol Caerfaddon.

18. Ochr yn ochr â'r rhaglen hon o drafodaethau bord gron a gweithdai technegol, gwnaethom ymgysylltu â 42 o randdeiliaid drwy ymgysylltu allanol lle y cyflwynwyd y fframwaith rheoleiddio deallusrwydd artiffisial a amlinellwyd yn y papur gwyn. Gwnaethom hefyd gynnal 28 o gyfarfodydd dwyochrog a chyfarfodydd â 16 o reoleiddwyr fel rhan o'n gwaith parhaus i gefnogi'r broses weithredu. Rydym yn cynnwys dirnadaethau o'r gweithgarwch ymgysylltu hwn yn yr ymateb i'r ymgynghoriad.

## Canfyddiadau ymgysylltu

19. Yn yr adran hon, rydym yn rhoi trosolwg cryno o'n digwyddiadau bord gron a'n gweithdai, sy'n crynhoi dirnadaethau yn bedwar maes yn seiliedig ar ffocws digwyddiadau bord gron a chyfranogiad:

- rheoleiddwyr.

- diwydiant.
- cymdeithas sifil.
- sefydliadau ymchwil.

## Rheoleiddwyr

20. Cynhaliwyd chwe digwyddiad bord gron â rheoleiddwyr er mwyn deall galluoedd ac anghenion presennol, gan gynnwys sut y byddai'r dull gweithredu a nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial yn cael ei roi ar waith mewn sectorau penodol, gan gynnwys iechyd a gofal cymdeithasol, cyfiawnder, addysg, a thrafnidiaeth.

21. Nododd y rheoleiddwyr lefelau amrywiol o wybodaeth a galluoedd mewnol ym maes deallusrwydd artiffisial, gyda'r rhan fwyaf yn cefnogi mesurau canolog i wella arbenigedd technegol. Cytunodd rhai y gallai cronfa o arbenigedd wella eu gallu i reoleiddio, ac awgrymodd eraill y gallai'r swyddogaeth ganolog arfaethedig roi arweiniad a deunyddiau hyfforddiant i reoleiddwyr.

22. Roedd y rheoleiddwyr yn gefnogol ar y cyfan i'r swyddogaeth ganolog a amlinellwyd yn y papur gwyn, gan bwysleisio y gallai fod yn bwynt cyswllt defnyddiol i reoleiddwyr. Fodd bynnag, pwysleisiodd y rheoleiddwyr na ddylai'r swyddogaeth ganolog amharu ar annibyniaeth na chylchoedd gwaith statudol presennol rheoleiddwyr, gan awgrymu na ddylai unrhyw ganllawiau i reoleiddwyr ar roi'r egwyddorion ar waith ymyrryd â mandadau na gwaith presennol rheoleiddwyr, eu dyblygu, na'u gwrthddweud.

23. Pwysleisiodd y rhai a gymerodd ran yn y digwyddiadau bord gron fod angen adnoddau digonol ar reoleiddwyr, ac roeddent yn cefnogi buddsoddiad y llywodraeth mewn galluoedd technegol. Nododd rhai ei bod yn bosibl y bydd angen i'r llywodraeth gyflwyno pwerau rheoleiddio newydd er mwyn i'r fframwaith fod yn effeithiol, gan nodi y bydd angen i'r llywodraeth, er mwyn sicrhau tryloywder a'r gallu i herio, o bosibl orfodi'r rhai sy'n datblygu ac yn cyflwyno deallusrwydd artiffisial i ddatgelu gwybodaeth ar adegau penodol.

24. Codwyd sawl her i oruchwyliaeth effeithiol gan reoleiddwyr yn ymwneud yn benodol â deallusrwydd artiffisial, gan gynnwys ffiniau swyddogaethol anhysbys a newidiol, diffyg tryloywder technegol, amgylcheddau anrhagweladwy, diffyg goruchwyliaeth neu fewnbwn dynol, a chylchoedd oes technolegol iteraidd iawn. Awgrymodd y rheoleiddwyr fod cydweithio rhwng rheoleiddwyr, peirianwyr diogelwch, ac arbenigwyr deallusrwydd artiffisial yn allweddol i greu mesurau dilysu cadarn sy'n atal, yn lleihau, ac yn lliniaru risgiau.

25. Er i'r rheoleiddwyr nodi bod yr egwyddorion yn cynnig tir cyffredin defnyddiol rhwng sectorau, nodwyd ganddynt y byddai angen dadansoddi sectorau penodol er mwyn dod o hyd i fylchau yn y fframwaith. Nododd rhai y byddai achosion defnydd sector-benodol yn helpu rheoleiddwyr i gymhwyso'r egwyddorion yn eu priod feysydd.

## Diwydiant

26. Clywsom gan amrywiaeth o randdeiliaid diwydiant yn y saith digwyddiad bord gron, gyda phynciau yn amrywio o'r gallu i ryngweithredu'n rhyngwladol, deallusrwydd artiffisial cyfrifol mewn diwydiant, deallusrwydd artiffisial diben cyffredinol, ac anghenion llywodraethu a safonau technegol.

27. Roedd rhai cyfranogwyr yn pryderu bod diffyg cydbwysedd yn y farchnad yn atal arloesedd a chystadleuaeth yn ecosystem deallusrwydd artiffisial. Yn benodol, roedd cyfranogwyr yn dadlau y byddai data mwy hygyrch, olrheiniadwy, ac atebol yn hyrwyddo arloesedd, gan nodi bod yn rhaid i gwmnïau bach yn aml ddibynnu ar arweinwyr

presennol yn y farchnad neu setiau data o ansawdd is oherwydd prinder setiau data perchenogol, masnachol fforddiadwy. Awgrymodd cyfranogwyr y byddai safonau clir ar gyfer data a mynediad tecach at setiau data o ansawdd uwch yn symbylu arloesedd mewn deallusrwydd artiffisial ar draws yr ecosystem ehangach ac atal manteision i'r rhai sydd eisoes yn rhan ohoni.

28. Nododd cyfranogwyr hefyd y gallai rhai o'r mesurau posibl i reoleiddio deallusrwydd artiffisial ei gwneud yn bosibl i arweinwyr presennol y farchnad ymwreiddio ymhellach ac ychwanegu at y diffyg cydbwysedd sydd eisoes yn bodoli yn y farchnad. Nododd cyfranogwyr y gallai busnesau bach a'r gymuned ffynhonnell agored wynebu baich cydymffurfio sylweddol, gyda rhai yn awgrymu y dylid defnyddio pyllau tywod rheoleiddio i brofi effaith rheoleiddio. Er bod rhai wedi awgrymu y dylai cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial gael ei neilltuo i gamau cynharach o'r cylch oes, roedd eraill yn rhybuddio y byddai gosod y cyfrifoldeb cyfreithiol am gymwysiadau i lawr y gadwyn ar ddatblygwyr ffynhonnell agored yn cyfyngu'n fawr ar arloesedd gan na fyddent yn gallu cyfrif am yr holl ddefnyddiau cod ffynhonnell agored posibl.

29. Nid oedd unrhyw gonsensws ynghylch a fyddai gofynion trwyddedu ar gyfer modelau sylfaen yn ffordd effeithiol o annog arloesi'n gyfrifol ym maes deallusrwydd artiffisial neu a fyddent, yn hytrach, yn canoli grym y farchnad yn nwylo ychydig o gwmnïau sefydledig. Nododd ychydig o'r cyfranogwyr y byddai canllawiau ymarferol ar weithredu ac achosion defnydd yn helpu sefydliadau i gymhwyso'r egwyddorion. Nododd rhai cyfranogwyr y gallai fframwaith trwyddedu a oedd ond yn caniatáu mynediad agored i rai rhannau o god system deallusrwydd artiffisial gadw rhai o fanteision rhannu gwybodaeth a thryloywder sy'n rhan annatod o ffynhonnell agored.

30. Nododd rhai cyfranogwyr nad yw'n glir pwy sy'n gyfrifol am reoleiddio deallusrwydd artiffisial, gan ddadlau o blaid rheoleiddiwr penodol, newydd ar gyfer deallusrwydd artiffisial neu reoleiddiwr arweiniol amlwg. Pwysleisiodd cyfranogwyr bwysigrwydd arbenigedd technegol i reoleiddio effeithiol.

31. Nododd cyfranogwyr hefyd bwysigrwydd rôl y gallu i ryngweithredu'n rhyngwladol, yswiriant, safonau technegol, a thryloywder i lwyddiant deallusrwydd artiffisial ar y farchnad.

## Cymdeithas sifil ac ymddiriedaeth y cyhoedd

32. Cynhaliwyd tri digwyddiad bord gron gyda busnesau bach, rhanddeiliaid cymdeithas sifil, a grwpiau buddiant arbennig i drafod ymddiriedaeth y cyhoedd ac effaith deallusrwydd artiffisial ar ddinasyddion a chymunedau wedi'u hymyleiddio.

33. Pwysleisiodd cyfranogwyr fod tegwch a chynhwysiant yn allweddol i wireddu manteision deallusrwydd artiffisial i bawb. Nododd cyfranogwyr bwysigrwydd amrywiaeth o ran y data a ddefnyddir i hyfforddi ac adeiladu deallusrwydd artiffisial, yn ogystal â'r timau sy'n datblygu, yn cyflwyno, ac yn rheoleiddio deallusrwydd artiffisial. Awgrymodd cyfranogwyr y byddai cyd-greu a chydweithio â chymunedau wedi'u hymyleiddio yn sicrhau y gallai deallusrwydd artiffisial greu manteision i bawb.

34. Pwysleisiodd cyfranogwyr fod angen i sefydliadau sy'n defnyddio deallusrwydd artiffisial nid yn unig bod yn dryloyw ynglŷn â phryd a sut y defnyddir deallusrwydd artiffisial ond y dylent hefyd sicrhau bod esboniadau yn hygyrch i grwpiau gwahanol. Nododd cyfranogwyr, er y gall deallusrwydd artiffisial gynnig manteision i gymunedau wedi'u hymyleiddio, fod y poblogaethau hyn yn aml yn wynebu effaith negyddol anghymesur o



ganlyniad i ddeallusrwydd artiffisial. Galwodd cyfranogwyr am fwy o addysg ar ddefnyddio deallusrwydd artiffisial ar y sail bod ymwybyddiaeth y cyhoedd, gwybodaeth sefydliadol, a llwybrau hygyrch i geisio iawn yn brin iawn ar hyn o bryd.

35. Nododd cyfranogwyr fod gan reoleiddwyr rôl allweddol i'w chwarae i'w gwneud yn haws i'r rhai y mae niweidiau sy'n gysylltiedig â deallusrwydd artiffisial yn effeithio arnynt herio'r canlyniadau hyn a cheisio iawn yn eu cylch. Pwysleisiodd cyfranogwyr fod angen cyllid ac adnoddau digonol ar reoleiddwyr er mwyn cyflawni hyn. Roedd cyfranogwyr yn cefnogi'n gryf y dylid cael ombwdsmon canolog ar gyfer deallusrwydd artiffisial er mwyn gwella hygyrchedd cyngor cyfreithiol o ansawdd uchel ar ddeallusrwydd artiffisial. Nododd llawer fod cyngor cyfreithiol ar ddeallusrwydd artiffisial yn gostus ar hyn o bryd, yn anodd cael gafael arno, ac weithiau'n cael ei roi gan ddarparwyr heb eu rheoleiddio y tu allan i'r proffesiwn cyfreithiol. Nododd cyfranogwyr hefyd y byddai'r ombwdsmon yn debygol o gael nifer mawr o gwynion bach, y dylai fod yn barod i ymdrin â nhw.

36. Roedd cyfranogwyr hefyd yn dadlau o blaid pwysigrwydd mesurau diogelu penodol i bobl ifanc, gan gynnwys newidiadau posibl i systemau statudol presennol megis y rhai ar gyfer diogelu data a chydraddoldeb.

### Academyddion, sefydliadau ymchwil, a melinau trafod

37. Cynhaliwyd tri digwyddiad i glywed gan academyddion, sefydliadau ymchwil, a melinau trafod ar ddiogelwch deallusrwydd artiffisial, cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial, a Chenhadaeth Technoleg Deallusrwydd Artiffisial UKRI.

38. Awgrymodd cyfranogwyr fod angen gwahaniaethu rhwng y mathau o risg sy'n codi o ganlyniad i ddeallusrwydd artiffisial, gan nodi y byddai angen i risgiau uniongyrchol a hirdymor gael eu hystyried wrth lunio unrhyw fesurau diogelwch ar gyfer deallusrwydd artiffisial. Roedd cyfranogwyr yn teimlo y dylai dadansoddiad yn ôl sector lywio asesiadau o risgiau sy'n gysylltiedig â deallusrwydd artiffisial. Nododd cyfranogwyr y gall diffyg tryloywder deallusrwydd artiffisial ei gwneud yn anodd i sefydliadau a rheoleiddwyr ganfod yr hyn sy'n achosi unrhyw niweidiau sy'n codi. Pwysleisiodd cyfranogwyr, er mwyn atal niweidiau, fod mesurau cyn cyflwyno yn allweddol i sicrhau bod deallusrwydd artiffisial yn ddiogel i'w ryddhau ar y farchnad.

39. Dadleuodd cyfranogwyr y gall trefniadau rheoleiddio o ansawdd uchel helpu deallusrwydd artiffisial i symud yn gyflym ac yn ddiogel o'r cam datblygu i'r farchnad. Dadleuodd cyfranogwyr fod angen mwy o wybodaeth dechnegol ymhob rhan o'r llywodraeth ac ymhlith rheoleiddwyr, ynghyd â sgiliau deallusrwydd artiffisial gwell drwy'r ecosystem ehangach gyfan. Galwodd rhai am ardystio peirianwyr deallusrwydd artiffisial a datblygwyr er mwyn ennyn mwy o ymddiriediaeth ymhlith y cyhoedd, a dadleuodd un arall o blaid ardystio swyddogion arweiniol sefydliadau sy'n gyfrifol am benderfyniadau sy'n ymwneud â deallusrwydd artiffisial. Nid oedd unrhyw gonsensws ynghylch ai rheoleiddiwr canolog, newydd ar gyfer deallusrwydd artiffisial ynteu'r rheoleiddwyr presennol fyddai'n rhoi'r fframwaith arfaethedig ar waith yn fwyaf effeithiol. Fodd bynnag, cytunodd cyfranogwyr y byddai cysoni canllawiau rheoleiddio a rhannu arbenigedd rhwng sectorau yn meithrin y gallu i gydymffurfio. Awgrymodd cyfranogwyr "economi gymysg" o reoleiddio, gyda gofynion statudol i sicrhau bod y rheolau'n gweithio'n effeithiol.

40. Nododd cyfranogwyr fod cylchoedd oes deallusrwydd artiffisial yn amrywiol ac yn gymhleth. Roedd cyfranogwyr am i'r llywodraeth ddiffinio pwy sy'n cymryd rhan ar adegau gwahanol o gylch oes deallusrwydd artiffisial a phennu rhwymedigaethau cyfatebol i egluro'r dirwedd. Fodd bynnag, nid oes unrhyw gytundeb ynglŷn â'r ffordd orau o wneud hyn, gyda chyfranogwyr yn awgrymu diffinio yn ôl swyddogaeth (fel yn achos rheoleiddio diogelu data), grym y farchnad neu fantais (fel yn achos rheoleiddio marchnadoedd digidol),

neu agosrwydd i risgiau a rhagweladwyedd rhesymol risgiau (fel yn achos deddfwriaeth diogelwch cynhyrchion). Er bod rhai cyfranogwyr am weld cyfrifoldebau llymach i ddatblygwyr modelau sylfaenol, roedd eraill yn rhybuddio y gallai ffocws rhy gul olygu y byddai cyfleoedd eraill sy'n gysylltiedig â deallusrwydd artiffisial yn cael eu colli o bosibl.

# Atodiad B: Rhestr o'r ymatebwyr i'r ymgynghoriad

## Rhestr o'r ymatebwyr i'r ymgynghoriad

1. Rydym yn ddiolchgar i'r holl unigolion a sefydliadau a rannodd eu dirnadaethau â ni yn ystod y cyfnod ymgynghori.
2. Bwriedir i'n fframwaith rheoleiddio deallusrwydd artiffisial fod yn gydweithredol a byddwn yn parhau i weithio'n agos gyda rheoleiddwyr, academyddion, cymdeithas sifil, a'r cyhoedd i fonitro a gwerthuso effeithiolrwydd ein dull gweithredu.
3. Yn unol â'n hysbysiad preifatrwydd<sup>123</sup> a chytundeb preifatrwydd yr arolwg ar-lein, dim ond enwau'r unigolion a'r sefydliadau hynny a gyflwynodd dystiolaeth drwy ein harolwg ar-lein ac a gydsyniodd i'n cytundeb preifatrwydd a restrir isod. Mae'r rhestr yn cynrychioli'r 209 o ymatebion a gyflwynwyd i'r arolwg ar-lein a ddadansoddwyd gennym ar ôl glanhau'r data i gael gwared ar ymatebion dyblyg, gwag, a sbam (gweler Atodiad A am fanylion). Rhestrir yr enwau fel y'u rhoddwyd. Dilëwyd enwau personol os oedd enw sefydliad ar gael. Rydym yn darparu 207 o enwau yma am fod dau ymateb yn ddienw.
4. Ceir rhagor o fanylion am y rhai a gyflwynodd ymatebion ysgrifenedig dros e-bost a thrwy'r arolwg ar-lein yn ôl y math o sefydliad a'r sector yn y dull estynedig o ddadansoddi ymatebion ysgrifenedig yn Atodiad A.

## Ymatebwyr i arolwg ar-lein yr ymgynghoriad

- |   |                                  |   |
|---|----------------------------------|---|
| 1. Adarga Limited   | 9. Alethics.ai                   | 19. Cyngor Celfyddydau Lloegr   |
| 2. ADS Group  | 10. Alacriter                    | 20. Pwyllgor Polisi Technoleg Cymdeithas Peiriannau Cyfrifiannu Ewrop |
| 3. Advai Ltd  | 11. Alacriter                    | 21. Cymdeithas Diwydiannau Technoleg Iechyd Prydain                   |
| 4. AGENCY: Assuring Citizen Agency in a World with Complex Online Harms       | 12. Aligned AI                   | 22. Cymdeithas y Cyfrifwyr Ardystiedig Siartredig                     |
| 5. Agile Property & Homes Limited   | 13. Cynghrair Eiddo Deallus      | 23. Cymdeithas y Mentrau Cydfuddiannol Ariannol                       |
| 6. AI & Partners  | 14. Altered Ltd                  | 24. Cymdeithas y Darlunwyr  |
| 7. Canolfan Deallusrwydd Artiffisial ar gyfer Gofal Iechyd seiliedig ar Werth | 15. Amendolara Holdings Limited  |   |
| 8. Aidan Freeman  | 16. Anton                        |   |
|   | 17. Arran McCutcheon             |   |
|   | 18. ART-AI, Prifysgol Caerfaddon |   |

<sup>123</sup> [Office for Artificial Intelligence – information collection and analysis: privacy notice](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

- |  |  |  |
|--|--|--|
| 25. Cyhoeddwr Cymdeithas Ddysgedig a Phroffesiynol   | 47. Cisco  | 76. Falcon Windsor   |
| 26. Assuring Autonomy International Programme, Prifysgol Caerefrog                                     | 48. Dinas, Prifysgol Llundain                          | 77. FlyingBinary   |
| 27. Avi Semelr   | 49. Cogstack   | 78. ForHumanity  |
| 28. Baringa Partners LLP   | 50. Colin Hayhurst                                     | 79. Freeths LLP  |
| 29. Barnacle Labs  | 51. Congenica Ltd                                      | 80. Fujitsu  |
| 30. Barry O'Brien  | 52. Craig MeulenCynghrair Hawliau Crewyr               | 81. Full Fact  |
| 31. Ben Hopkinson  | 53. CTRL-Shift & Collider Health                       | 82. Geeks Ltd.   |
| 32. BPI British Diwydiant Ffonograffig   | 54. Cyferd   | 83. Getty Images   |
| 33. Bristows LLP   | 55. CyLon Ventures                                     | 84. GlaxoSmithKline plc  |
| 34. Cyngor Hawlfraint Prydain  | 56. DACS (Cymdeithas Hawlfraint Dylunio ac Artistiaid) | 85. Glenn Donaldson  |
| 35. Cymdeithas Rheoli Plâu Prydain   | 57. Daniel Marsden                                     | 86. Global Witness   |
| 36. Cymdeithas Diwydiant Diogelwch Prydain   | 58. Darrell Warner Limited                             | 87. Greg Colbourn  |
| 37. Canolfan Deallusrwydd Artifisial: Datblygiadau Cymdeithasol ac Arloesol Prifysgol Brunel Llundain. | 59. Deborah W.A. Foulkes                               | 88. Greg Mathews   |
| 38. Grŵp BSI, Yr Iseldiroedd   | 60. Deloitte UK  | 89. Guy Warren   |
| 39. Grŵp BT  | 61. Cynghrair y Datblygwyr                             | 90. Hazy   |
| 40. Bud Financial  | 62. Adran Addysg                                       | 91. Henry  |
| 41. Calvin Karpenko  | 63. Direct Line Group                                  | 92. Hollie   |
| 42. Carlo Attubato   | 64. DNV  | 93. Hugging Face   |
| 43. Canolfan Polisi Deallusrwydd Artifisial a Digidol Washington, DC. UDA                              | 65. Dr. Michael K. Cohen                               | 94. Iain Darby   |
| 44. Canolfan Astudiaethau Polisi   | 66. easyJet Airline Company Ltd.                       | 95. IFPI   |
| 45. Charlie Bowler   | 67. Ed Hagger  | 96. INRO London  |
| 46. Chegg, Inc.  | 68. EKC Group  | 97. Sefydliad Dyfodol Gwaith   |
|  | 69. Elliott Andrews                                    | 98. Y Sefydliad Cyfrifwyr Siartredig yng Nghymru a Lloegr                                  |
|  | 70. Emily Gray   | 99. Sefydliad Arloesi a Chyfnawid Gwybodaeth   |
|  | 71. Emma Ahmed-Rengers                                 | 100. Sefydliad Ffiseg a Pheirianeg mewn Meddygaeth   |
|  | 72. Enzai Technologies Limited                         | 101. Sefydliad Ffiseg a Pheirianeg mewn Meddygaeth (Grŵp Cyfrifiannu Clinigol a Gwyddonol) |
|  | 73. Equity   |  |
|  | 74. Eviden   |  |
|  | 75. Experian UK&I                                      |  |

- |  |   |   |
|--|---|---|
| 102. Sefydliad Diogelwch ac Iechyd Galwedigaethol                    | 126. Martin Gore  | 154. Cymdeithas y Cyhoeddwyr Proffesiynol           |
| 103. Ffederasiwn Rhyngwladol y Newyddiadurwyr                        | 127. Mastercard Europe  | 155. Yr Athro Julia Black                           |
| 104. Jake Bailey   | 128. MedTech Europe   | 156. PRS for Music                                  |
| 105. Jake Wilkinson  | 129. Megha Barot  | 157. Cymdeithas y Cyhoeddwyr                        |
| 106. Cymdeithas Diwydiannau Electroneg a Thechnoleg Gwybodaeth Japan | 130. Michael Fisher   | 158. Gwasanaethau Trwyddedu'r Cyhoeddwyr            |
| 107. Joe Collman   | 131. Michael Pascu  | 159. Pupils 2 Parliament                            |
| 108. Johnny Luk  | 132. Microsoft  | 160. Queen Bee Marketing Hive                       |
| 109. Johnson & Johnson   | 133. Mind Foundry   | 161. Rebecca Palmer                                 |
| 110. Jonas Herold-Zanker   | 134. Mukesh Sharma  | 162. RELX   |
| 111. Joseph Johnston   | 135. Labordy Ffisegol Cenedlaethol                                      | 163. Reset  |
| 112. Judith Barker   | 136. Sefydliad Undeb Trethdalwyr Cenedlaethol                           | 164. Rohan Vij                                      |
| 113. Kainos Software Ltd   | 137. Cymdeithas Undeb y Newyddiadurwyr                                  | 165. Cymdeithas Ffotograffig Frenhinol Prydain Fawr |
| 114. Kelechi Ejikeme   | 138. NATS   | 166. Salesforce                                     |
| 115. Knowledge Associates Cambridge Ltd.                             | 139. Nebuli Ltd.  | 167. SambaNova Systems inc                          |
| 116. Labour for the Long Term  | 140. Prifysgol Newcastle  | 168. Samuel Frewin                                  |
| 117. Legal & General Group PLC                                       | 141. Newsstand  | 169. SAP  |
| 118. Canolfan Dyfodol Deallusrwydd Leverhumle                        | 142. Nicole Hawkesford  | 170. Scale AI                                       |
| 119. Lewis   | 143. Swyddfa Safonau mewn Addysg, Gwasanaethau Plant a Sgiliau (Ofsted) | 171. ScaleUp Institute                              |
| 120. LSE Law, Technology and Society Group                           | 144. Y Swyddfa Rheoleiddio Ystadegau                                    | 172. Scott Timcke                                   |
| 121. Lucy Purdon   | 145. Orbit RRI  | 173. Seldon   |
| 122. Luke Richards   | 146. Paul Dunn  | 174. Sharon Darcy                                   |
| 123. Lumi Network  | 147. Paul Evans   | 175. Simon Kirby                                    |
| 124. Y Gymdeithas Ymchwil i'r Farchnad                               | 148. Paul Ratcliffe   | 176. Skin Analytics Ltd                             |
| 125. Marta   | 149. Pearson  | 177. South West Grid for Learning                   |
|  | 150. Phrasee  | 178. Stability AI                                   |
|  | 151. Pippa Robertson  | 179. Steve Kendall                                  |
|  | 152. Planar AI Limited  | 180. STFC Hartree Centre                            |
|  | 153. Policy Connect   |   |

- |  |  |   |
|--|--|---|
| 181. Surrey Institute for People-Centred Artificial Intelligence   | 188. Cymdeithas Ymchwil Weithredol     | 200. Valentino Giudice  |
| 182. Teal Legal Ltd  | 189. Cymdeithas Amddiffyn y Fferyllwyr | 201. ValidMind  |
| 183. Siambrau Temple Garden  | 190. Y Gymdeithas Seicolegol           | 202. W Legal Ltd  |
| 184. Asiantaeth Trwyddedu Hawlfraint Cyf   | 191. Cymdeithas y Cyhoeddwyr           | 203. Rhwydwaith Cymunedau Mwy Diogel Cymru (sy'n cynnwys yr Heddlu, Tân, Awdurdodau Lleol, y Gwasanaeth Prawf a'r Trydydd Sector), a gynhelir gan CLILC |
| 185. Canolfan Arloesi'r Labordy Data   | 192. Cymdeithas yr Awduron             | 204. Cyngor Sir Swydd Warwick   |
| 186. Sefydliad Gwasanaeth Cwsmeriaid   | 193. Prifysgol Caer-wynt               | 205. We and AI  |
| 187. Gwasanaeth cynghori amlasiantaethol (MASS) ar Ddeallusrwydd Artiffisial a Digidol Gwasanaeth Rheoliadau ar gyfer iechyd a gofal cymdeithasol, | 194. Tom Edward Ashworth               | 206. Workday  |
|  | 195. TRANSEARCH International          | 207. Urdd Llenorion Prydain Fawr  |
|  | 196. Trilateral Research               |   |
|  | 197. Prifysgol Caeredin                |   |
|  | 198. Prifysgol Caeredin                |   |
|  | 199. Prifysgol Caer-wynt               |   |

# Atodiad C: Crynodebau o gwestiynau unigol

## Yr egwyddorion deallusrwydd artiffisial traws-sector diwygiedig

### 1. A ydych chi'n cytuno y byddai gofyniad ar sefydliadau i'w gwneud yn glir eu bod yn defnyddio deallusrwydd artiffisial yn sicrhau gwell tryloywder?

1. Roedd y mwyafrif o'r ymatebwyr yn cytuno bod gofyniad ar sefydliadau i'w gwneud yn glir eu bod yn defnyddio deallusrwydd artiffisial yn ddigon i sicrhau tryloywder. Roedd yr ymatebwyr a anghytunodd naill ai'n teimlo y byddai labelu'r defnydd o ddeallusrwydd artiffisial yn annigonol neu'n faich anghymesur.

2. Nododd ymatebwyr a ddadleuodd y byddai'r mesur yn annigonol yn aml nad oes gan reoleiddwyr y pwerau perthnasol, y cyllid, na'r galluoedd i sicrhau tryloywder yn ddigonol. Yn gysylltiedig â hyn, nododd ymatebwyr broblemau o ran gorfodi a mynediad at apeliadau ac iawn. Argymhellodd rhai ymatebwyr y dylai'r llywodraeth ystyried mesurau statudol a systemau atebolrwydd perthnasol. Awgrymodd ychydig o'r ymatebwyr y dylai esboniadau gael eu targedu at y cyd-destun a'r gynulleidfa.

3. Roedd ymatebwyr eraill yn pryderu y byddai gofyniad cyffredinol ynglŷn â thryloywder yn creu rhwystr beichus i gymwysiadau deallusrwydd artiffisial risg is. Nododd yn ymatebydd fod y cynnig yn tybio y bydd gan gyfranogwr unigol yn y gadwyn werth deallusrwydd artiffisial drosolwg digonol dros y camau niferus o bosibl mewn cylchoedd oes a chymwysiadau. Roedd ychydig o'r ymatebwyr am weld trothwyon clir (gan gynnwys "cymwysiadau risg uchel") a chanllawiau gan y llywodraeth a'r rheoleiddwyr ar ofynion ynglŷn â thryloywder.

4. Roedd ymatebwyr yn pryderu y gall mesurau tryloywder ryngweithio o bosibl â deddfwriaeth gyfredol ac arfaethedig, megis diogelu data ac eiddo deallusol.

### 2. A oes unrhyw fesurau eraill y gallem eu gosod ar sefydliadau i wella tryloywder o ran deallusrwydd artiffisial?

5. Cafwyd cefnogaeth gref i nifer o fesurau tryloywder ymhlith ymatebwyr. Pwysleisiodd ymatebwyr fod tryloywder yn allweddol i ennyn ymddiriedaeth y cyhoedd, atebolrwydd, a fframwaith rheoleiddio effeithiol a dilysadwy.

6. Roedd llawer o'r ymatebwyr yn cymeradwyo rhwymedigaethau adrodd clir ar y mewnbynnau a ddefnyddir i adeiladu a hyfforddi deallusrwydd artiffisial. Nododd ymatebwyr y gellid gwella tryloywder drwy ddatgelu nifer o fewnbynnau, o ddata i gyfrifiannu. Gan adlewyrchu ymatebion i gwestiwn F1 ynglŷn â modelau sylfaen, cafwyd pryderon ynghylch a oedd ansawdd data hyfforddi yn ddigon da, a oedd yn cydymffurfio â fframweithiau cyfreithiol cyfredol, gan gynnwys eiddo deallusol a diogelu data, ac a oedd yn briodol ar gyfer defnyddiau i lawr y gadwyn. Roedd ychydig o'r ymatebwyr yn dadlau y byddai datgelu cyfrifiannu yn sicrhau gwell tryloywder ar effeithiau amgylcheddol deallusrwydd artiffisial.

7. Roedd llawer o'r ymatebwyr hefyd yn cefnogi labelu'r defnydd o ddeallusrwydd artiffisial ac allbynnau deallusrwydd artiffisial, gyda llawer yn argymhell y mesur i wella ymwybyddiaeth defnyddwyr ac atebolrwydd sefydliadol. Awgrymodd rhai ymatebwyr y byddai labelu allbynnau a gynhyrchir drwy ddeallusrwydd artiffisial yn helpu i wrthsefyll

camwybodaeth a gynhyrchir drwy deallusrwydd artiffisial a hyrwyddo hawliau eiddo deallusol. Roedd ychydig o'r ymatebwyr am weld dewisiadau optio i mewn cliriach ar gyfer defnyddiau o ddata a deallusrwydd artiffisial, gydag opsiynau ar gyfer dewisiadau amgen dynol.

8. Roedd rhai ymatebwyr yn cymeradwyo mesurau a fyddai'n annog esboniadau ar gyfer canlyniadau ac effeithiau posibl deallusrwydd artiffisial. Mae hyn yn cynnwys mesurau sy'n dangos i ddefnyddwyr sut roedd modelau yn cynhyrchu allbynnau neu atebion sydd hefyd yn ymdrin â chyfyngiadau modelau a'u heffeithiau. Yn yr un modd, nododd ychydig o'r ymatebwyr bwysigrwydd addysg sefydliadau ac addysg i'r cyhoedd drwy wybodaeth hygyrch ac ymdrechion wedi'u targedu i godi ymwybyddiaeth. Awgrymodd un neu ddau o'r ymatebwyr y byddai cofrestrau cyhoeddus neu sefydliadol ar gyfer deallusrwydd artiffisial (risg uchel) yn helpu i godi ymwybyddiaeth.

9. Er bod rhai ymatebwyr yn dadlau o blaid adrodd ar fanylion modelau, roedd llawer yn pwysleisio mai datgelu gwybodaeth dechnegol gymhleth i reoleiddwyr a dilyswyr annibynnol, yn hytrach na'r cyhoedd, fyddai'r opsiwn gorau. Awgrymodd ymatebwyr y dylai sefydliadau rannu manylion technegol modelau megis pwysau, paramedrau, defnyddiau, a phrofion. Nododd ymatebwyr y dylai asesiadau o'r effaith ac asesiadau risg, yn ogystal â phenderfyniadau llywodraethu a marchnata, fod ar gael naill ai i reoleiddwyr neu'r cyhoedd, gydag ychydig yn nodi cyfaddawdu posibl o ran cyfrinachau masnach. Roedd rhai ymatebwyr yn cymeradwyo technegau sicrwydd annibynnol, megis archwiliadau trydydd parti a safonau technegol.

10. Awgrymodd ychydig o'r ymatebwyr y dylid egluro hawliau a chyfrifoldebau cyfreithiol sy'n gysylltiedig â deallusrwydd artiffisial, gydag ychydig o'r rhain yn argymhell y dylid cyflwyno deddfwriaeth a mesurau diffyg cydymffurfio ynglŷn â deallusrwydd artiffisial.

### **3. A ydych chi'n cytuno bod y llwybrau presennol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch yn ddigonol?**

11. Nododd dros hanner yr ymatebwyr nad oedd y llwybrau presennol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch drwy'r fframweithiau cyfreithiol presennol yn ddigonol. Yn benodol, nododd ymatebwyr fod diffyg tryloywder ynglŷn â phryd a sut y defnyddir deallusrwydd artiffisial yn atal defnyddwyr rhag gallu nodi niweidiau sy'n gysylltiedig â deallusrwydd artiffisial. Yn yr un modd, nododd ymatebwyr fod diffyg tryloywder ynglŷn â'r data a ddefnyddir i hyfforddi modelau deallusrwydd artiffisial yn cymhlethu mesurau diogelu data ac yn atal deiliaid hawliau eiddo deallusol rhag arfer eu hawliau cyfreithiol a moesol. Nododd ychydig o'r ymatebwyr gostau uchel ymgyfreitha unigol hefyd, gan ddadlau o blaid llwybrau cliriach i weithredu'n unigol ac ar y cyd.

### **4. A allai'r llwybrau presennol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch gael eu gwella, os o gwbl.**

12. Roedd llawer o'r ymatebwyr am weld y llywodraeth yn egluro hawliau a chyfrifoldebau cyfreithiol sy'n gysylltiedig â deallusrwydd artiffisial, ond nid oedd unrhyw gonsensws ynglŷn â sut y dylid gwneud hyn. Awgrymodd llawer o'r ymatebwyr y dylai hawliau a chyfrifoldebau mewn deddfwriaeth bresennol gael eu hegluro drwy ddulliau megis canllawiau rheoleiddiol. Roedd ymatebwyr hefyd yn awyddus yn gyffredinol i weld trefn ganolog ar ffurfiau gwahanol, gyda rhai ymatebwyr yn dadlau o blaid creu dull canolog o geisio iawn megis rheoleiddiwr deallusrwydd artiffisial canolog, corff goruchwyllo, swyddogaeth gydgyssylltu, neu rheoleiddiwr arweiniol. Roedd rhai ymatebwyr am weld gofynion statudol pellach, megis trwyddedu.



13. Pwysleisiodd llawer o'r ymatebwyr bwysigrwydd tryloywder ystyrion a phwysleisiodd rhai fod angen llwybrau hygyrch i geisio iawn. Roedd ymatebwyr yn teimlo y byddai mesurau i ddangos i ddefnyddwyr pryd a sut y defnyddir deallusrwydd artiffisial yn helpu unigolion i nodi pryd a sut roedd niweidiau wedi digwydd. Roedd ymatebwyr am weld llwybrau clir – ac mewn rhai achosion, gorfodol – i herio penderfyniadau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch. Nododd ymatebwyr broblemau gydag ymglyfreitha costus, yn enwedig mewn perthynas ag ymyrryd â hawliau eiddo deallusol. Roedd ymatebwyr yn teimlo y byddai mwy o dryloywder ar gyfer systemau deallusrwydd artiffisial yn ei gwneud yn haws i geisio iawn am ystod eang o niweidiau posibl ac, yn yr un modd, y byddai egluro llwybrau i geisio iawn yn sicrhau gwell tryloywder. Nododd rhai ymatebwyr bwysigrwydd cytundebau rhyngwladol i sicrhau llwybrau effeithiol i herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial dros ffiniau neu geisio iawn yn eu cylch. Dim ond ychydig o'r ymatebwyr a gododd fesurau megis moratoria a switsys diffodd gorfodol.

#### **5. A ydych chi'n cytuno y bydd yr egwyddorion traws-sector diwygiedig, unwaith y cânt eu rhoi ar waith yn effeithiol, yn ymdrin â'r risgiau sy'n codi o ganlyniad i dechnolegau deallusrwydd artiffisial?**

14. Roedd y mwyafrif o'r ymatebwyr yn cytuno y byddai'r egwyddorion yn ymdrin â'r risgiau sy'n codi o ganlyniad i dechnolegau deallusrwydd artiffisial pan fyddant wedi'u rhoi ar waith yn effeithiol. Roedd yr ymatebwyr a anghytunodd yn tueddu i gyfeirio at bryderon ynglŷn â gorfodi a diffyg cefnogaeth statudol i'r egwyddorion neu faterion ehangach ynglŷn â pharodrwydd rheoleiddwyr, gan gynnwys adnoddau, galluoedd, a chydgyssylltu.

15. Nododd ymatebwyr yn aml fod angen i'r fframwaith fod yn hyblyg, canolbwyntio ar gyd-destun, a chael ei ategu gan weithgarwch monitro a gwerthuso, gan gyfeirio at ba mor gyflym roedd technoleg yn newid.

16. Roedd ychydig o'r ymatebwyr yn teimlo nad oedd y cwestiwn yn glir a gofynnwyd am ragor o fanylion ynglŷn â gweithredu effeithiol.

#### **6. Beth, os unrhyw beth, sydd ar goll yn yr egwyddorion diwygiedig?**

17. Dadleuodd llawer o'r ymatebwyr y dylai'r egwyddorion deallusrwydd artiffisial traws-sector gynnwys hawliau dynol a ffyniant dynol yn fwy penodol, gan nodi y dylai deallusrwydd artiffisial gael ei ddefnyddio i wella bywyd pobl. Roedd ymatebwyr yn dadlau o blaid gwahanol hawliau dynol a gwerthoedd cysylltiedig, gan gynnwys rhyddid, plwraliaeth, preifatrwydd, cydraddoldeb, cynhwysiant, a hygyrchedd.

18. Roedd rhai ymatebwyr am weld rhagor o fanylion ynglŷn â rhoi'r egwyddorion ar waith. Roedd yr ymatebwyr hyn yn aml yn gofyn am fwy o fanylion am adnoddau rheoleiddwyr, gan nodi y byddai angen i reoleiddwyr gael adnoddau, sgiliau, a phwerau digonol er mwyn rhoi'r egwyddorion ar waith "yn effeithiol". Gofynnodd un neu ddau o'r ymatebwyr am fwy o eglurder ynglŷn â sut y disgwylir i reoleiddwyr a sefydliadau ymdrin â chyfaddawdu, megis rhwng esboniadwyedd a chywirdeb a thryloywder a phreifatrwydd.

19. Yn gysylltiedig â hyn, roedd rhai ymatebwyr am gael rhagor o arweiniad ynglŷn â sut y byddai'r egwyddorion deallusrwydd artiffisial yn rhyngweithio â deddfwriaeth bresennol ac yn cael eu rhoi ar waith drwyddynt. Cododd ymatebwyr bryderon o ran cyfraith diogelu data ac eiddo deallusol yn bennaf, ond gofynnodd ychydig o'r ymatebwyr am gael ymdeimlad mwy holistaidd o'r ffordd y bydd y llywodraeth yn ymdrin â deallusrwydd artiffisial o ran strategaethau adrannol, megis strategaeth deallusrwydd artiffisial y Weinyddiaeth Amddiffyn. Nododd rhai ymatebwyr y byddai'r egwyddorion yn aneffeithiol heb statws statudol, gydag ychydig yn pwysleisio pwysigrwydd gorfodi systemau atebolrwydd am ddeallusrwydd artiffisial.

20. Roedd rhai ymatebwyr yn dadlau y dylai'r egwyddorion ymdrin â nifer o faterion sy'n gysylltiedig â chadernid gweithredol. Yn yr ymatebion hyn awgrymwyd mesurau ar gyfer diogelwch a seiberddiogelwch digonol, prosesau datgomisiynu, diogelu cystadleurwydd, sicrhau mynediad, ac ymdrin â risgiau sy'n gysylltiedig â gorddibyniaeth. Roedd nifer tebyg o ymatebwyr am weld egwyddorion penodol ynglŷn ag ansawdd data a chysondeb rhyngwladol.

21. Argymhellodd ychydig o'r ymatebwyr y byddai cynnwys egwyddorion yn amlwg yn cyfateb i risgiau systemig ac effeithiau cymdeithasol ehangach, cynaliadwyedd, neu addysg a llythrennedd. O ran risgiau systemig, roedd ymatebwyr yn tueddu i godi pryderon ynglŷn â'r niweidiau posibl y gall technolegau deallusrwydd artiffisial eu hachosi i ddemocratiaeth a rheolaeth y gyfraith o ran twyllwybodaeth ac ymyrryd mewn etholiadau.

## Dyletswydd statudol i roi sylw

### **7. A ydych chi'n cytuno y byddai cyflwyno dyletswydd statudol ar reoleiddwyr i roi sylw priodol i'r egwyddorion yn egluro ac yn cryfhau mandadau rheoleiddwyr i roi ein hegwyddorion ar waith, a chadw ymagwedd hyblyg tuag at eu rhoi ar waith?**

22. Roedd dros hanner yr ymatebwyr yn cytuno i ryw raddau neu'n cytuno'n gryf y byddai dyletswydd statudol yn egluro ac yn cryfhau mandad rheoleiddwyr i roi'r fframwaith ar waith. Fodd bynnag, cododd llawer rybuddion a nodir yn fanwl yng Nghwestiwn 8.

### **8. A oes ymyriad statudol amgen a fyddai'n fwy effeithiol?**

23. Roedd llawer yn teimlo y byddai mesurau statudol wedi'u targedu, gan gynnwys pwerau estynedig i reoleiddwyr, yn ymyriad statudol mwy effeithiol. Yn benodol, nododd ymatebwyr fod angen i reoleiddwyr gael pwerau ymchwilio priodol. Roedd rhai hefyd am weld canlyniadau diffyg cydymffurfiaeth yn cael eu diffinio'n gliriach. Awgrymodd ymatebwyr hefyd y byddai deddfwriaeth benodol ynglŷn â deallusrwydd artiffisial, rheoleiddiwr deallusrwydd artiffisial newydd, a rheolau llym ynglŷn â defnyddio deallusrwydd artiffisial mewn rhai cyd-destunau yn ymyriadau statudol mwy effeithiol. Soniodd un neu ddau o'r ymatebwyr y dylai unrhyw ddyletswyddau o ran deallusrwydd artiffisial gael eu gosod ar y rhai sy'n gweithredu yn y farchnad, yn hytrach na'r rheoleiddwyr.

24. Roedd rhai ymatebwyr yn teimlo mai'r ddyletswydd statudol arfaethedig yw'r ymyriad mwyaf effeithiol ac y dylid ei gweithredu. Fodd bynnag, gwnaeth ymatebwyr eraill, wrth fynegi eu cefnogaeth, godi pryderon ehangach na fyddai'r fframwaith yn ddigon gorfodadwy heb ryw fath o statws statudol. Pwysleisiodd bron chwarter yr ymatebwyr y byddai angen adnoddau a galluooedd ychwanegol ar reoleiddwyr er mwyn iddynt weithredu dyletswydd statudol yn effeithiol. Roedd ymatebwyr eraill yn teimlo y gallai dyletswydd i roi sylw darfu ar y gyfundrefn reoleiddio, arloesedd ac ymddiriedaeth pe bai'n cael ei gweithredu'n fryslog. Argymhellodd yr ymatebwyr hyn y dylai'r ddyletswydd gael ei hadolygu ar ôl cyfnod o weithredu anstatudol, yn enwedig i weld a oedd unrhyw ryngweithio â chyfreithiau a chylchoedd gwaith rheoleiddio sy'n bodoli eisoes. Nododd ychydig o'r ymatebwyr nad oedd nod terfynol na therfynau amser y fframwaith rheoleiddio deallusrwydd artiffisial yn ddigon clir, a oedd yn achosi ansicrwydd.

25. Roedd rhywfaint o gefnogaeth o blaid y Llywodraeth yn gorfodi mesurau megis archwiliadau gan drydydd parti, ardystio, a mesurau cadwyni cyflenwi yn debyg i fesurau Amgylcheddol, Cymdeithasol a Llywodraethu, gan gynnwys adroddiadau ar ddata hyfforddi. Roedd ychydig o'r ymatebwyr yn cefnogi system fonitro ganolog i olrhain cydymffurfiaeth reoleiddiol a thechnolegau newydd a all olygu bod angen ehangu cwmpas rheoleiddio.

## Swyddogaethau canolog newydd i gefnogi'r fframwaith

### **9. A ydych chi'n cytuno y byddai'r swyddogaeth a amlinellwyd yn adran 3.3.1 yn fuddiol o ran ein fframwaith rheoleiddio deallusrwydd artiffisial pe bai'n cael ei chyflawni'n ganolog?**

26. Roedd bron pob un o'r ymatebwyr yn cytuno y byddai'n fuddiol i'r fframwaith pe bai'r swyddogaethau arfaethedig yn cael eu cyflawni'n ganolog, gyda llawer yn dadlau y byddai gweithgareddau canolog yn ei gwneud yn bosibl i'r llywodraeth fonitro ac iteru'r fframwaith. Awgrymodd llawer y dylai adborth gan reoleiddwyr, diwydiant, y byd academaidd, cymdeithas sifil, a'r cyhoedd gael ei ddefnyddio i fesur effeithiolrwydd, gyda rhai'n galw am bwytiau adolygu rheolaidd i asesu a oedd y swyddogaeth ganolog yn dal yn addas at y diben. Roedd ychydig o'r ymatebwyr yn pryderu bod rhai o'r gweithgareddau arfaethedig eisoes yn cael eu cyflawni gan sefydliadau eraill, gan awgrymu y dylid mapio'r gwaith presennol er mwyn osgoi dyblygu.

### **10. Beth, os unrhyw beth, sydd ar goll yn swyddogaethau canolog?**

27. Er bod ymatebwyr yn cefnogi'r swyddogaethau canolog arfaethedig yn gyffredinol, roedd llawer am weld mwy o fanylion am gyflawni pob gweithgaredd, gyda rhai ymatebwyr yn cefnogi pwyslais cryfach ar ymgysylltu a phartneriaethau â sefydliadau presennol.

28. Nododd ymatebion bwysigrwydd ymdrin â risgiau sy'n gysylltiedig â deallusrwydd artiffisial ac ennyn ymddiriedaeth y cyhoedd mewn technolegau deallusrwydd artiffisial. Awgrymodd rhai ymatebwyr y dylai'r llywodraeth flaenoriaethu'r swyddogaeth arfaethedig o ran risgiau, gan nodi pa mor bwysig ydoedd bod risgiau sy'n gysylltiedig â deallusrwydd artiffisial yn cael eu nodi a'u hasesu. Nododd ymatebwyr y dylai'r dadansoddiad risg hwn gynnwys risgiau moesegol, megis tuedd, a risgiau systemig i gymdeithas, megis newidiadau i'r farchnad lafur. Pwysleisiodd ychydig o'r ymatebwyr y byddai'r swyddogaeth o ran addysg ac ymwybyddiaeth yn allweddol i ennyn ymddiriedaeth y cyhoedd.

29. Nododd ymatebwyr bwysigrwydd cysondeb rheoleiddiol rhwng sectorau a chyfundrefnau rhyngwladol. Roedd rhai ymatebwyr yn dadlau y dylai'r swyddogaethau canolog gynnwys mwy ar y gallu i ryngweithredu, gan nodi bod seiberddiogelwch, twyllwbyodaeth, a thorri hawlfraint yn faterion y bydd angen cydweithio rhyngwladol yn eu cylch.

30. Awgrymodd rhai ymatebwyr y dylai fod statws statudol i rai neu bob un o'r swyddogaethau canolog neu y dylent gael eu cyflawni gan gorff annibynnol. Pwysleisiodd ymatebwyr hefyd y dylai'r swyddogaethau canolog, er mwyn bod yn effeithiol, gael adnoddau digonol a'r arbenigedd technegol angenrheidiol. Nodwyd bod hyn yn arbennig o bwysig o ran mapio risgiau, sganio'r gorwel, a swyddogaethau monitro a gwerthuso.

31. Roedd gweithgareddau neu swyddogaethau ychwanegol a awgrymwyd gan ymatebwyr yn cynnwys: pwerau statudol i sicrhau diogelwch modelau deallusrwydd artiffisial tra galluog; cydgysylltu â'r gweinyddiaethau datganoledig; a goruchwylio cydymffurfiaeth deallusrwydd artiffisial â chyfreithiau presennol, gan gynnwys fframweithiau eiddo deallusol a diogelu data.

### **11. A wyddoch am unrhyw sefydliadau sydd eisoes yn bodoli a ddylai gyflawni un neu fwy o'n swyddogaethau canolog arfaethedig?**

32. Ar y cyfan, roedd tua chwarter yr ymatebwyr yn teimlo y dylai'r llywodraeth gyflawni un neu fwy o'r swyddogaethau canolog. Nododd ymatebwyr sefydliadau eraill hefyd a allai gefnogi'r swyddogaethau canolog, gan gynnwys rheoleiddwyr, sefydliadau sy'n ymchwilio i dechnolegau a melinau trafod, cwmnïau yn y sector preifat, a grwpiau ymchwil

academaidd. Roedd llawer o'r ymatebwyr yn dadlau o blaid adeiladu'r swyddogaethau rheoleiddio ar gryfderau presennol ecosystem reoleiddio'r DU. Nododd ymatebwyr y gallai mentrau cydgysylltu rheoleiddio fel y Fforwm Cydweithredu Rheoleiddio Digidol helpu i nodi bylchau yng nghylchoedd gwaith rheoleiddwyr ac ymateb iddynt. Nododd ymatebwyr hefyd fod melinau trafod a sefydliadau ymchwil megis Sefydliad Alan Turing, Sefydliad Ada Lovelace, a Sefydliad Dyfodol Gwaith yn ymgymryd neu wedi ymgymryd â gweithgareddau a all ategu'r rhai a ddisgrifir yn y swyddogaethau canolog arfaethedig.

## **12. A oes gweithgareddau ychwanegol a fyddai'n helpu busnesau i arloesi a defnyddio technolegau deallusrwydd artiffisial yn hyderus?**

33. Roedd llawer o'r ymatebwyr yn teimlo y gallai'r swyddogaethau canolog gyflawni gweithgareddau pellach i helpu busnesau i gymhwysu'r egwyddorion at arferion pob dydd sy'n gysylltiedig â deallusrwydd artiffisial. Roedd ymatebwyr yn dadlau y dylai'r llywodraeth a rheoleiddwyr roi cymorth i ddiwydiant ar ffurf rhaglenni hyfforddiant ac adnoddau addysgol. Nododd ymatebwyr y byddai'r cymorth hwn yn arbennig o bwysig i sefydliadau sy'n gweithredu ar draws neu rhwng sectorau.

34. Roedd ymatebwyr yn teimlo y dylai rheoleiddwyr ddatblygu canllawiau i alluogi busnesau i arloesi'n hyderus a'u diweddarau'n rheolaidd. Nododd ymatebwyr y gallai prosesau cydymffurfio digyswllt a chostus lesteirio arloesedd ac arafu'r broses o fabwysiadu deallusrwydd artiffisial.

35. Awgrymodd ymatebwyr y gallai'r llywodraeth wella mynediad at ddata o ansawdd uchel, sicrhau cysondeb rhyngwladol ynglŷn â gofynion deallusrwydd artiffisial, a hwyluso cydweithio rhwng rheoleiddwyr, diwydiant, a'r byd academaidd. Nododd rhai ymatebwyr bod arloesi'n gyfrifol ym maes deallusrwydd artiffisial yn cael ei ategu gan fynediad at ddata amrywiol, ansawdd uchel, sy'n dod o ffynonellau moesegol. Awgrymodd ymatebwyr y gallai ymddiriedolaethau data a noddur gan y llywodraeth helpu i wella mynediad at ddata. Roedd rhai ymatebwyr o'r farn y gallai'r llywodraeth chwarae rôl allweddol i sicrhau bod cyfundrefnau rheoleiddio deallusrwydd artiffisial yn cael eu cysoni'n rhyngwladol, gan nodi y byddai'r gallu i ryngweithredu yn hyrwyddo masnach a chystadleuaeth. Awgrymodd ychydig o'r ymatebwyr y gallai'r llywodraeth hwyluso cydweithio rhwng rheoleiddwyr, diwydiant, a'r byd academaidd er mwyn sicrhau cysondeb rhwng rheoleiddio deallusrwydd artiffisial, arloesi, ac ymchwil. Awgrymodd nifer bach o ymatebwyr y dylid cyflwyno deddfwriaeth ynglŷn â deallusrwydd artiffisial yn hytrach na swyddogaethau canolog er mwyn rhoi mwy o sicrwydd cyfreithiol.

### **12.1. Os felly, a ddylai'r gweithgareddau hyn gael eu cyflawni gan y llywodraeth, rheoleiddwyr, neu sefydliad gwahanol?**

36. Er bod ymatebwyr wedi nodi rhai gweithgareddau i helpu busnesau i arloesi a defnyddio technolegau deallusrwydd artiffisial yn hyderus a ddylai gael eu harwain gan reoleiddwyr, awgrymodd mwyafrif o'r ymatebwyr mai'r llywodraeth a ddylai gyflawni'r gweithgareddau hyn.

## **13. A oes gweithgareddau ychwanegol a fyddai'n helpu unigolion a defnyddwyr i ddefnyddio technolegau deallusrwydd artiffisial yn hyderus?**

37. Roedd ymatebwyr yn rhoi blaenoriaeth ar dryloywder o blith yr egwyddorion traws-sector, gyda bron hanner yn dadlau y dylai fod modd i unigolion a defnyddwyr nodi pryd a sut mae gwasanaeth neu sefydliad yn defnyddio deallusrwydd artiffisial.

38. Roedd llawer o'r ymatebwyr yn teimlo y byddai addysg a hyfforddiant yn ennyn ymddiriedaeth y cyhoedd mewn technolegau deallusrwydd artiffisial a helpu i gyflymu'r broses o'u mabwysiadu. Pwysleisiodd ymatebwyr y dylid gwella llythrennedd deallusrwydd

artiffisial drwy addysg a hyfforddiant sy'n galluogi defnyddwyr i ddefnyddio cynhyrchion a gwasanaethau deallusrwydd artiffisial yn fwy effeithiol. Awgrymodd ymatebwyr y dylai hyfforddiant gwmpasu pob cam o gylch oes deallusrwydd artiffisial a meithrin dealltwriaeth o fanteision deallusrwydd artiffisial yn ogystal â'i risgiau. Nododd ymatebwyr y dylai sefydliadau addysg, defnyddwyr ac eirioli, ochr yn ochr â'r llywodraeth a'r rheoleiddwyr, helpu i sicrhau bod gwybodaeth yn hygyrch.

39. Roedd rhai ymatebwyr am weld llwybrau cliriach i ddefnyddwyr herio niweidiau sy'n gysylltiedig â deallusrwydd artiffisial neu geisio iawn yn eu cylch. Pwysleisiodd rhai bwysigrwydd mesurau diogelu data digonol. Nododd ychydig o'r ymatebwyr y byddai deddfwriaeth benodol ynglŷn â deallusrwydd artiffisial yn rhoi sicrwydd cyfreithiol ac yn helpu i ennyn ymddiriedaeth y cyhoedd.

### **13.1 Os felly, a ddylai'r gweithgareddau hyn gael eu cyflawni gan y llywodraeth, rheoleiddwyr, neu sefydliad gwahanol?**

40. Er bod y rhan fwyaf o'r ymatebwyr wedi argymhell y dylai'r llywodraeth, rheoleiddwyr, diwydiant, a chymdeithas sifil weithio gyda'i gilydd i helpu unigolion a defnyddwyr i ddefnyddio technolegau deallusrwydd artiffisial yn hyderus, awgrymodd bron hanner yr ymatebwyr mai'r llywodraeth a ddylai ymgymryd â gweithgareddau i wella hyder defnyddwyr mewn deallusrwydd artiffisial.

### **14. Sut y gallwn osgoi sefyllfa lle mae canllawiau ar ddeallusrwydd artiffisial a gyhoeddir gan wahanol reoleiddwyr yn gorgyffwrdd, yn dyblygu, neu'n gwrthddweud ei gilydd?**

41. Awgrymodd llawer o'r ymatebwyr mai'r swyddogaethau canolog arfaethedig fyddai'r dull mwyaf effeithiol o osgoi canllawiau sy'n gorgyffwrdd, yn dyblygu, neu'n gwrthddweud ei gilydd. Nododd ymatebwyr y byddai'r swyddogaethau canolog yn helpu rheoleiddwyr drwy nodi risgiau traws-sector, hwyluso camau i reoli risg yn gyson, rhoi arweiniad ar faterion traws-sector, a monitro a gwerthuso'r fframwaith yn ei gyfanrwydd.

42. Er bod ymatebwyr wedi pwysleisio y byddai angen cydgysylltu rheoleiddiol er mwyn sicrhau bod y fframwaith yn cael ei roi ar waith yn gyson rhwng cylchoedd gwaith, nid oedd unrhyw gytundeb ar y ffordd orau o gyflawni hyn. Awgrymodd rhai y dylid sefydlu rheoleiddiwr deallusrwydd artiffisial newydd, cynigiodd ychydig y dylid penodi un o'r rheoleiddwyr presennol yn 'rheoleiddiwr arweiniol', ac roedd eraill yn cefnogi mesurau cydgysylltu rheoleiddiol gwirfoddol, gan bwysleisio rôl fforymau rheoleiddio megis y Fforwm Cydweithredu Rheoleiddio Digidol.

43. Awgrymodd rhai ymatebwyr y byddai safonau a thechnegau sicrwydd traws-sector llorweddol yn hyrwyddo cysondeb rhwng cylchoedd gwaith rheoleiddwyr, sectorau, ac awdurdodaethau rhyngwladol. Argymhellodd ymatebwyr y dylid egluro cylchoedd gwaith penodol pob rheoleiddiwr mewn perthynas â deallusrwydd artiffisial er mwyn hyrwyddo cydlyniant yn y dirwedd reoleiddio. Roedd ychydig yn dadlau y byddai cyflwyno deddfwriaeth ynglŷn â deallusrwydd artiffisial, gan gynnwys gwneud yr egwyddorion deallusrwydd artiffisial a'r trefniadau cydgysylltu rheoleiddiol yn statudol, yn atal gwahaniaethau rhwng rheoleiddwyr.

## Monitro a gwerthuso'r fframwaith

### 15. A ydych chi'n cytuno â'n dull gweithredu cyffredinol o fonitro a gwerthuso?

44. Roedd dros hanner yr ymatebwyr yn cytuno â'r dull cyffredinol o fonitro a gwerthuso a nodwyd yn y papur gwyn ar reoleiddio deallusrwydd artiffisial. Roedd llawer yn canmol y cynigion ar gyfer dolen adborth, gan gynghori y dylid defnyddio diwydiant, rheoleiddwyr, a chymdeithas sifil i helpu i fesur effeithiolrwydd y fframwaith. Roedd ymatebwyr yn cefnogi dull iteraidd ar y cyfan, ac awgrymodd rhai y dylid ymgynghori â diwydiant fel rhan o werthusiad rheolaidd i asesu ac addasu'r fframwaith. Roedd ychydig o'r ymatebwyr yn dadlau o blaid cyhoeddi canfyddiadau gwerthusiadau o'r fframwaith.

45. Nododd rhai ymatebwyr nad oedd digon o fanylion neu fod y dull o fonitro a gwerthuso yn aneglur. Er mwyn cadarnhau pa mor ymarferol yw'r dull hwn, gofynnodd rhai ymatebwyr am ragor o wybodaeth am fformat, mynychder, a ffynonellau'r data a fydd yn cael eu datblygu a'u defnyddio. Pwysleisiodd rhai o'r ymatebwyr hyn bwysigrwydd nodi problemau gyda'r fframwaith mewn modd amserol. Pwysleisiodd ymatebwyr y bydd angen i risgiau deallusrwydd artiffisial gael eu monitro'n barhaus, gan nodi y bydd angen mwy o eglurder a thryloywder ynglŷn â sut y bydd risgiau yn cael eu huwchgwyfeirio a'u trin.

### 16. Beth yw'r ffordd orau o fesur effaith ein fframwaith?

46. Awgrymodd llawer o'r ymatebwyr mai dull gweithredu a lywir gan ddata fyddai'r ffordd fwyaf effeithiol o fesur effaith fframwaith. Argymhellodd ymatebwyr ymarferion casglu data ansoddol a meintiol, asesiadau o'r effaith, a dangosyddion perfformiad allweddol (DPAau). Roedd enghreifftiau o DPAau posibl yn cynnwys ymddiriedaeth a boddhad defnyddwyr, cyfradd arloesi, amser i gyrraedd y farchnad, cwynion a digwyddiadau andwyol, ymglyfreitha, a chostau cydymffurfiaeth. Awgrymodd ychydig o'r ymatebwyr y dylid defnyddio twf economaidd i fesur effaith y fframwaith. Roedd un neu ddau am weld mesuriadau wedi'u teilwra at sectorau penodol, gan awgrymu y dylai'r llywodraeth ymgysylltu â rheoleiddwyr er mwyn deall sut maent yn mesur effeithiau rheoleiddio ar eu priod ddiwydiannau.

47. Argymhellodd ychydig dros chwarter yr ymatebwyr y dylai'r llywodraeth gynnal deialog agos â diwydiant, cymdeithas sifil, a phartneriaid rhyngwladol. Pwysleisiodd ymatebwyr dro ar ôl tro pa mor bwysig ydoedd bod barn holistaidd yn cael ei chasglu ar effaith, gyda llawer yn nodi y dylai'r llywodraeth ymgysylltu â rhanddeiliaid sy'n gallu cynnig safbwyntiau gwahanol ar effeithlonrwydd y fframwaith, gan gynnwys egin fusnesau a busnesau bach. Roedd ymatebwyr yn teimlo y byddai ymgynghoriad eang i gasglu tystiolaeth o agweddau'r cyhoedd tuag at y fframwaith a deallusrwydd artiffisial yn fwy cyffredinol hefyd yn ddefnyddiol.

48. Awgrymodd ymatebwyr y dylai'r gallu i ryngweithredu'n rhyngwladol gael ei fonitro er mwyn sicrhau bod y fframwaith yn ei gwneud yn bosibl i fusnesau fasnachu â marchnadoedd rhyngwladol a datblygu cynhyrchion ar eu cyfer. Awgrymodd rhai ymatebwyr y dylid cyfeirio at ddangosyddion a fframweithiau sefydledig, megis Nodau Datblygu Cynaliadwy'r Cenhedloedd Unedig a'r Pum Cyfalaf, er mwyn llywio set o fesurau ansoddol a meintiol.

**17 A ydych chi'n cytuno bod ein dull gweithredu yn taro'r cydbwysedd cywir rhwng cefnogi arloesedd ym maes deallusrwydd artiffisial; ymdrin â risgiau hysbys, wedi'u blaenoriaethu; a diogelu'r fframwaith rheoleiddio deallusrwydd artiffisial rhag datblygiadau yn y dyfodol?**

49. Roedd hanner yr ymatebwyr yn cytuno bod y dull gweithredu yn taro'r cydbwysedd cywir rhwng cefnogi arloesedd ym maes deallusrwydd artiffisial; ymdrin â risgiau hysbys, wedi'u blaenoriaethu; a diogelu'r fframwaith rheoleiddio deallusrwydd artiffisial rhag datblygiadau yn y dyfodol. Fodd bynnag, roedd rhai ymatebwyr yn pryderu na fyddai'r dull gweithredu yn gallu datblygu ochr yn ochr â datblygiadau technolegol ym maes deallusrwydd artiffisial, gan nodi y bydd y gallu i ddiogelu'r fframwaith yn ddigonol rhag datblygiadau yn y dyfodol yn dibynnu ar gadw hyblygrwydd a'r gallu i addasu wrth roi'r egwyddorion ar waith. Roedd ymatebwyr am gael rhagor o eglurder ynglŷn â'r meysydd penodol sydd i'w rheoleiddio, gan bwysleisio bod angen i reoleiddwyr fod yn rhagweithiol o ran nodi'r risg o niwed.

50. Roedd dros draean o'r ymatebwyr yn anghytuno. Roedd ymatebwyr yn pryderu nad yw'r fframwaith yn neilltuo cyfrifoldeb am ganlyniadau deallusrwydd artiffisial yn glir. Roedd rhai yn meddwl y byddai'r ffocws ar arloesi ym maes deallusrwydd artiffisial, twf economaidd, a chreu swyddi yn atal ffocws digonol ar risgiau sy'n gysylltiedig â deallusrwydd artiffisial, megis tuedd a gwahaniaethu.

## Galluoedd rheoleiddwyr

**18. A ydych chi'n cytuno mai rheoleiddwyr sydd yn y sefyllfa orau i gymhwyso'r egwyddorion ac mai'r llywodraeth sydd yn y sefyllfa orau i oruchwylio a chyflawni swyddogaethau canolog?**

51. Roedd bron pob un o'r ymatebwyr yn cytuno mai rheoleiddwyr sydd yn y sefyllfa orau i roi'r egwyddorion ar waith ac mai'r llywodraeth sydd yn y sefyllfa orau i oruchwylio a chyflawni'r swyddogaethau canolog.

52. Er bod ymatebwyr wedi nodi y byddai arbenigedd rheoleiddwyr ar eu priod feysydd yn allweddol i deilwra'r egwyddorion traws-sector at anghenion sectorau yn effeithiol, awgrymodd rhai hefyd y dylai'r llywodraeth helpu rheoleiddwyr i reoli risgiau deallusrwydd artiffisial o fewn eu cylchoedd gwaith drwy feithrin eu sgiliau a'u harbenigedd ym maes deallusrwydd artiffisial.

53. Roedd rhai ymatebwyr yn dadlau y byddai angen i'r llywodraeth weithio'n agos gyda'r rheoleiddwyr i oruchwylio'r fframwaith yn effeithiol a chyflawni'r swyddogaethau canolog. Roedd rhai hefyd yn cefnogi rhagor o gydweithio rhwng rheoleiddwyr. Roedd ychydig yn teimlo y dylai proses y llywodraeth o oruchwylio'r fframwaith fod yn agored ac yn dryloyw, gan ddadlau o blaid mewnbwn gan ddiwydiant a chymdeithas sifil.

54. Roedd rhai ymatebwyr yn pryderu nad oedd yr un corff presennol yn y sefyllfa orau i helpu i roi'r fframwaith arfaethedig ar waith a'i oruchwylio, gydag ychydig yn gofyn am ddeddfwriaeth a rheoleiddiwr newydd ar gyfer deallusrwydd artiffisial.

**19. Fel rheoleiddiwr, pa gymorth fyddai ei angen arnoch i gymhwyso'r egwyddorion mewn ffordd gymesur sydd o blaid arloesi?**

55. Er bod y rheoleiddwyr a ymatebodd i'r cwestiwn hwn yn cefnogi'r fframwaith arfaethedig, roedd ychydig dros chwarter yn dadlau mai cydgysylltu fyddai'r brif her i weithredu mewn ffordd gymesur sydd o blaid arloesi. Roedd rheoleiddwyr o'r farn y byddai rhannu arferion gorau yn werthfawr er mwyn sicrhau cysondeb ac ymgorffori gwybodaeth

bresennol mewn dulliau gweithredu ar gyfer sectorau penodol. Awgrymodd llawer y byddai systemau cadarn ar gyfer rhannu gwybodaeth rhwng rheoleiddwyr a'r swyddogaethau canolog arfaethedig yn helpu i osgoi dyblygu gofynion rhwng nifer o reoleiddwyr.

56. Nododd y rheoleiddwyr a ymatebodd i'r cwestiwn hwn fod galluoedd anghyson ym maes deallusrwydd artiffisial, gyda thros chwarter yn gofyn am ragor o gymorth o ran arbenigedd technegol ac eraill yn dangos dulliau datblygedig o ymdrin â deallusrwydd artiffisial o fewn eu cylchoedd gwaith. Nododd rheoleiddwyr fylchau cyffredin mewn galluoedd, gan gynnwys diffyg gwybodaeth dechnegol am ddeallusrwydd artiffisial a dealltwriaeth gyfyngedig o ble a sut y defnyddir deallusrwydd artiffisial ei ddefnyddio gan y rhai y maent yn eu rheoleiddio. Awgrymodd rhai y byddai cymorth gan y llywodraeth i feithrin gallu sefydliadol mewnol yn eu helpu i gymhwyso'r egwyddorion yn effeithiol o fewn eu cylchoedd gwaith presennol, gyda rhai yn nodi eu bod yn ei chael hi'n anodd cystadlu â'r sector cyhoeddus i recriwtio'r arbenigedd a'r sgiliau technegol cywir. Nododd un neu ddau o'r rheoleiddwyr sut roedd mentrau megis Cronfa Arloeswyr y Rheoleiddwyr a ariennir gan y llywodraeth eisoes wedi ei gwneud yn bosibl iddynt ddatblygu dulliau o ymdrin ag arloesi'n gyfrifol ym maes deallusrwydd artiffisial yn eu cylchoedd gwaith. Nododd dau reoleiddwr fod cwmpas eu cylchoedd gwaith a'u pwerau statudol presennol mewn perthynas â deallusrwydd artiffisial yn aneglur. Gofynnodd y rheoleiddwyr hyn am ragor o fanylion ynglŷn â sut y byddai'r swyddogaeth ganolog yn sicrhau bod rheoleiddwyr yn defnyddio eu pwerau a'u cylchoedd gwaith mewn ffordd gydlynol wrth iddynt gymhwyso'r egwyddorion.

## **20. A ydych chi'n cytuno mai tîm cyfunol o arbenigwyr ar ddeallusrwydd artiffisial fyddai'r ffordd fwyaf effeithiol o lenwi bylchau mewn galluoedd a helpu rheoleiddwyr i gymhwyso'r egwyddorion?**

57. Roedd dros dri chwarter yr ymatebwyr yn cytuno mai tîm cyfunol o arbenigwyr ar ddeallusrwydd artiffisial fyddai'r ffordd fwyaf effeithiol o feithrin gallu cyffredin a llenwi bylchau. Roedd ymatebwyr yn teimlo y gallai tîm cyfunol o arbenigwyr ar ddeallusrwydd artiffisial helpu rheoleiddwyr i ddeall deallusrwydd artiffisial ac ymdrin â'i nodweddion unigryw o fewn eu sectorau, gan eu helpu i gymhwyso'r egwyddorion yn gyson rhwng cylchoedd gwaith.

58. Er bod ymatebwyr o blaid cynyddu mynediad rheoleiddwyr at arbenigedd ym maes deallusrwydd artiffisial, pwysleisiodd llawer y byddai angen i dîm cyfunol gynnwys safbwyntiau amrywiol ac amlddisgyblaethol. Roedd ymatebwyr yn teimlo y dylai'r tîm ddwyn ynghyd arbenigedd technegol ym maes deallusrwydd artiffisial a gwybodaeth am sectorau penodol, arbenigwyr diwydiant, a chymdeithas sifil er mwyn sicrhau bod rheoleiddwyr yn ystyried amrywiaeth eang o safbwyntiau wrth gymhwyso'r egwyddorion.

59. Nododd rhai ymatebwyr y byddai cronfa o arbenigwyr yn annigonol, gan awgrymu y dylid rhoi blaenoriaeth ar allu mewnol rheoleiddwyr ag arbenigedd ar sectorau penodol.

## **Adnoddau ar gyfer deallusrwydd artiffisial dibynadwy**

### **21. Pa adnoddau anrheoleiddiol ar gyfer deallusrwydd artiffisial dibynadwy fyddai fwyaf defnyddiol i helpu sefydliadau i ymgorffori egwyddorion rheoleiddio deallusrwydd artiffisial mewn prosesau busnes sy'n bodoli eisoes?**

60. Cafwyd cefnogaeth gref o blaid defnyddio safonau technegol a thechnegau sicrwydd. Roedd ymatebwyr yn cytuno y byddai'r naill a'r llall yn helpu sefydliadau i ymgorffori'r egwyddorion deallusrwydd artiffisial mewn prosesau busnes sy'n bodoli eisoes. Roedd llawer o'r ymatebwyr yn canmol gwaith Hwb Safonau Deallusrwydd Artiffisial y DUa'r



Ganolfan Moeseg Data ac Arloesi ar sicrwydd deallusrwydd artiffisial. Er bod rhai ymatebwyr wedi nodi y byddai llai o faich cydymffurfio ar fusnesau pe bai adnoddau a phrosesau yn gyson rhwng sectorau, nododd eraill bwysigrwydd adnoddau a phrosesau ychwanegol ar gyfer sectorau penodol. Awgrymodd ymatebwyr hefyd y dylid ategu safonau technegol ag astudiaethau achos ac enghreifftiau o arferion da.

61. Dadleuodd ymatebwyr y byddai adnoddau a thechnegau safonol ar gyfer nodi a lliniaru risgiau posibl sy'n gysylltiedig â deallusrwydd artiffisial hefyd yn helpu sefydliadau i ymgorffori'r egwyddorion deallusrwydd artiffisial. Nododd rhai fod technegau sicrwydd megis asesiadau o'r effaith ac asesiadau risg, monitro perfformiad modelau, gwerthuso ansicrwydd modelau, ac ymarferion tîm coch yn arbenigol o ddefnyddiol o ran nodi risgiau deallusrwydd artiffisial. Argymhellodd ychydig o'r ymatebwyr dechnegau sicrwydd y gellir eu defnyddio i ganfod ac atal problemau megis gadael i bethau lithro o ran lliniaru risgiau sy'n ymwneud â data. Er iddynt ganmol rôl adnoddau ar gyfer deallusrwydd artiffisial dibynadwy, nododd ychydig o'r ymatebwyr eu bod hefyd yn awyddus i weld mesurau rheoleiddio llymach, megis gofynion statudol ar gyfer cymwysiadau risg uchel o ddeallusrwydd artiffisial neu gorff gwarchod ar gyfer modelau sylfaen.

62. Roedd ymatebwyr yn teimlo y gall adnoddau a thechnegau megis metrigau tegwch, adroddiadau tryloywder, a chanllawiau moeseg deallusrwydd artiffisial sefydliadol gefnogi'r defnydd cyfrifol o ddeallusrwydd artiffisial a hefyd feithrin ymddiriedaeth y cyhoedd yn y dechnoleg ar yr un pryd. Nododd ymatebwyr eu bod yn awyddus i fodolau deallusrwydd artiffisial gael eu dilysu gan drydydd parti drwy archwiliadau o duedd, cynlluniau labelu i ddefnyddwyr, ac ardystio allanol yn erbyn safonau technegol.

63. Nododd ychydig o'r ymatebwyr fanteision i sefydliadau a defnyddwyr fel ei gilydd pe bai dulliau o lywodraethau deallusrwydd artiffisial yn cael eu cysoni'n rhyngwladol. Roedd rhai yn cefnogi safonau technegol rhyngweithredu ar gyfer deallusrwydd artiffisial, gan ganmol sefydliadau datblygu safonau rhyngwladol megis y Sefydliad Safoni Rhyngwladol (ISO) a Sefydliad y Peirianwyr Trydanol ac Electroneg. Nododd eraill gryfder ystod o waith rhyngwladol ar ddeallusrwydd artiffisial, gan gynnwys gwaith gan wledydd unigol, megis Fframwaith Rheoli Risgiau Deallusrwydd Artiffisial Sefydliad Safonau a Thechnoleg Cenedlaethol (NIST) UDA a Sefydliad Dilysu Deallusrwydd Artiffisial Singapore, ochr yn ochr â gwaith ar lywodraethu rhyngwladol gan gyrff amlochrog megis y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, y Cenhedloedd Unedig, a fforwm y G7.

## Syniadau i gloi

### **22. A oes gennych unrhyw syniadau eraill ynglŷn â'n dull gweithredu cyffredinol? Nodwch unrhyw gyfleoedd coll, diffygion, a bylchau yn ein fframwaith.**

64. Roedd rhai ymatebwyr yn teimlo y byddai'n fuddiol pe bai'r fframwaith rheoleiddio deallusrwydd artiffisial a nodwyd yn y papur gwyn yn cynnwys canllawiau manylach ar risgiau sy'n gysylltiedig â deallusrwydd artiffisial. Roedd rhai yn awyddus i weld mesurau llymach ar gyfer risgiau difrifol, yn enwedig o ran y defnydd o ddeallusrwydd artiffisial mewn cyd-destun lle mae diogelwch yn hanfodol. Awgrymodd ymatebwyr y byddai'r fframwaith yn gliriach pe bai'r llywodraeth yn rhoi categorïau o risg ar gyfer defnyddiau penodol o ddeallusrwydd artiffisial megis gorfodi'r gyfraith a gweithleoedd. Pwysleisiodd ymatebwyr eraill y gall deallusrwydd artiffisial achosi neu gyflymu risgiau sylweddol ynglŷn â phreifatrwydd a thorri cyfraith diogelu data, seiberymosodiadau, ymyrryd mewn etholiadau, camwybodaeth, ymyrryd â hawliau dynol, cynaliadwyedd amgylcheddol, a materion sy'n ymwneud â chystadleuaeth. Roedd ychydig o'r ymatebwyr yn pryderu am y risg ddifodol bosibl sy'n codi o ganlyniad i ddeallusrwydd artiffisial. Roedd llawer o'r ymatebwyr yn teimlo bod technolegau deallusrwydd artiffisial yn datblygu'n gynt na'r prosesau rheoleiddio.

65. Roedd rhai ymatebwyr yn dadlau bod llwyddiant y fframwaith yn dibynnu ar ddigon o gydgyssylltu rhwng rheoleiddwyr er mwyn cynnig dull clir a chyson o ymdrin â deallusrwydd artiffisial rhwng sectorau a marchnadoedd. Nododd ymatebwyr hefyd fod sectorau gwahanol yn wynebu manteision a risgiau penodol mewn cysylltiad â deallusrwydd artiffisial, gan awgrymu y byddai angen i'r fframwaith gydbwysu'r cysondeb a gynigir drwy ofynion traws-sector â chywirdeb dulliau gweithredu ar gyfer sectorau penodol. Yn benodol, nododd ymatebwyr y dylai unrhyw reolau neu gyrrff newydd i reoleiddio deallusrwydd artiffisial gael eu hadeiladu ar gyloedd gwaith statudol presennol rheoleiddwyr a safonau rheoleiddio perthnasol. Nododd ymatebwyr hefyd y byddai angen adnoddau digonol ar reoleiddwyr ar ffurf arbenigedd a sgiliau technegol i roi'r fframwaith ar waith yn effeithiol.

66. Pwysleisiodd ymatebwyr yn rheolaidd bwysigrwydd cysoni rhyngwladol i reoleiddio deallusrwydd artiffisial yn effeithiol. Awgrymodd ymatebwyr y dylai'r DU weithio tuag at ecosystem reoleiddio sy'n gyson yn rhyngwladol ar gyfer deallusrwydd artiffisial drwy ddatblygu fframwaith safon aur a hyrwyddo arferion gorau drwy sianeli amlochrog allweddol megis y Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd, y CU, fforwm y G7, a fforwm y G20. Nododd ymatebwyr y byddai dulliau o reoleiddio deallusrwydd artiffisial sy'n anghyson neu sy'n gorgyffwrdd yn achosi beichiau cydymffurfio sylweddol. Roedd ymatebwyr yn dadlau y gall cydweithredu rhyngwladol gefnogi arloesi'n gyfrifol ym maes deallusrwydd artiffisial yn y DU drwy greu rheolau clir a phendant sy'n ei gwneud yn bosibl i fuddsoddiadau symud ar draws sawl marchnad. Awgrymodd ymatebwyr hefyd y dylid sefydlu gweithgorau dwyochrog â phartneriaid strategol allweddol er mwyn rhannu arbenigedd. Pwysleisiodd rhai ymatebwyr y dylai dull gweithredu'r DU o blaid arloesi gael ei gyflwyno'n ddi-oed er mwyn aros yn gystadleuol mewn tirwedd ryngwladol sy'n newid yn gyflym.

## Cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial

### **L1. Pa heriau a allai godi pan fydd rheoleiddwyr yn cymhwyso'r egwyddorion mewn gwahanol gymwysiadau a systemau deallusrwydd artiffisial? Sut y gallem ymateb i'r heriau hyn drwy ein fframwaith rheoleiddio deallusrwydd artiffisial arfaethedig?**

67. Roedd ymatebwyr yn teimlo bod dwy her greiddiol i reoleiddwyr wrth gymhwyso'r egwyddorion rhwng cymwysiadau a systemau deallusrwydd artiffisial gwahanol: diffyg cyfrifoldeb cyfreithiol clir yn ystod cylchoedd oes deallusrwydd artiffisial cymhleth a phroblemau o ran cydgysylltu rhwng rheoleiddwyr a sectorau.

68. Roedd dros chwarter yr ymatebwyr yn teimlo nad oedd yn glir pwy fyddai'n cael ei ddwyn yn gyfrifol am risgiau sy'n gysylltiedig â deallusrwydd artiffisial. Cododd rhai ymatebwyr bryder pellach ynglŷn â rhyngweithio dryslyd rhwng y fframwaith a deddfwriaeth bresennol.

69. Er bod bron hanner yr ymatebwyr yn pryderu ynglŷn â chydgyssylltu a chysondeb rhwng sectorau a chylchoedd gwaith rheoleiddio, nododd rhai fod ateb i'w gael mewn dull gweithredu sy'n seiliedig ar gyd-destun (a fyddai'n cryfhau'r fframwaith). Gofynnodd ymatebwyr am ganllawiau gan reoleiddwyr sy'n seiliedig ar sectorau, adnoddau cydymffurfio, ac am i reoleiddwyr ymgysylltu â diwydiant.

70. Awgrymodd llawer o'r ymatebwyr y dylid cyflwyno gofynion statudol neu ganoli'r fframwaith mewn un corff sefydliadol, ond nid oedd unrhyw gonsensws ynghylch a ddylai canoli o'r fath fod ar ffurf rheoleiddiwr arweiniol, rheoleiddiwr canolog, neu swyddogaeth gydgyssylltu. Awgrymodd rhai ymatebwyr y dylid gwneud tryloywder diwydiant neu archwiliadau gan drydydd parti yn ofynnol.

71. Nododd ymatebwyr hefyd fod diffyg safonau a chytundebau rhyngwladol yn her, gan gyfeirio at bwysigrwydd cysondeb a chydweithio rhyngwladol.

**L2.i. A ydych chi'n cytuno, drwy roi ein hegwyddorion ar waith drwy fframweithiau cyfreithiol sy'n bodoli eisoes, y byddwn yn neilltuo cyfrifoldeb am ddeallusrwydd artiffisial drwy gydol cylch oes mewn ffordd deg ac effeithiol?**

72. Er bod rhai o'r ymatebwyr yn cytuno i ryw raddau y byddai'r egwyddorion yn neilltuo cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial mewn ffordd deg ac effeithiol drwy fframweithiau cyfreithiol sy'n bodoli eisoes, nid oedd y rhan fwyaf o'r ymatebwyr yn anghytuno neu nid oeddent yn cytuno nac yn anghytuno. Nododd llawer o'r ymatebwyr nad yw'n glir sut y byddai'r egwyddorion rheoleiddio deallusrwydd artiffisial yn cael eu rhoi ar waith drwy fframweithiau cyfreithiol sy'n bodoli eisoes. Mynegodd ymatebwyr bryderon ynglŷn â bylchau mewn deddfwriaeth bresennol, gan gynnwys eiddo deallusol, gwasanaethau cyfreithiol, a chyfraith cyflogaeth. Nododd rhai ymatebwyr fod angen i hawliau eiddo deallusol gael eu cadarnhau a'u hegluro er mwyn gwella cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial. Nododd ychydig o'r ymatebwyr fod angen i'r fframwaith deallusrwydd artiffisial fonitro ac ymaddasu wrth i'r dechnoleg ddatblygu a chael ei defnyddio'n fwy cyffredinol. Nododd un ymatebydd mai ar lefel cyflwynwyr y mae'r baich atebolrwydd ac awgrymodd y byddai'n hanfodol mynd i'r afael â bylchau yn y wybodaeth am gylch oes deallusrwydd artiffisial er mwyn gwella'r ffordd y neilltuir cyfrifoldeb cyfreithiol.

**L.2.ii. Sut y gellid ei wella, os o gwbl?**

73. Roedd llawer o'r ymatebwyr yn teimlo bod angen i'r fframwaith egluro atebolrwydd yn ystod cylch oes deallusrwydd artiffisial ymhellach. Yn benodol, nododd ymatebwyr dro ar ôl tro fod angen person a oedd yn gyfreithiol gyfrifol am ddeallusrwydd artiffisial ac awgrymodd rhai fodel a oedd yn debyg i Swyddogion Diogelu Data.

74. Nododd dros chwarter yr ymatebwyr y byddai angen deddfwriaeth neu bwerau newydd i reoleiddwyr ym maes deallusrwydd artiffisial er mwyn neilltuo atebolrwydd yn ystod y cylch oes yn effeithiol. Enwodd rhai fesurau penodol y byddai angen eu hategu'n statudol a dadleuodd rhai o blaid trwyddedu a chymeradwyo ymlaen llawn ac awgrymodd un neu ddau foratoriwm ar y deallusrwydd artiffisial lefel uchaf.

75. Roedd eraill yn teimlo mai'r ateb gorau oedd gwneud cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial yn fwy eglur yn unol â'r fframweithiau sydd eisoes yn bodoli. Roedd ymatebwyr am weld eglurder ynglŷn â sut y byddai'r egwyddorion yn cael eu cymhwyso gyda neu drwy gyfraith bresennol, gyda rhai yn awgrymu y byddai canllawiau rheoleiddiol yn rhoi mwy o sicrwydd.

76. Awgrymodd ymatebwyr hefyd y byddai mesurau anstatudol megis gwella galluoedd technegol rheoleiddwyr, safonau domestig a rhyngwladol, a thechnegau sicrwydd yn helpu i neilltuo cyfrifoldeb cyfreithiol yn ystod cylch oes deallusrwydd artiffisial mewn ffordd deg ac effeithiol.

77. Nododd eraill y byddai'r swyddogaethau canolog arfaethedig, gan gynnwys asesu risg, sganio'r gorwel, a monitro a gwerthuso, yn allweddol i sicrhau bod cyfrifoldeb cyfreithiol am ddeallusrwydd artiffisial yn cael ei ddosbarthu dros y cylch oes mewn ffordd deg ac effeithiol wrth i alluoedd deallusrwydd artiffisial ddatblygu a chael eu defnyddio fwyfwy.

### **L3. Os ydych yn gweithio i fusnes sy'n datblygu, yn defnyddio, neu'n gwerthu deallusrwydd artiffisial, sut rydych yn rheoli risg deallusrwydd artiffisial ar hyn o bryd, gan gynnwys drwy'r gadwyn gyflenwi ehangach? Sut y gallai'r llywodraeth gefnogi ffordd effeithiol o reoli risgiau sy'n gysylltiedig â deallusrwydd artiffisial?**

78. Dywedodd bron hanner yr ymatebwyr i'r cwestiwn hwn wrthym eu bod wedi rhoi prosesau asesu risg ar waith yn eu sefydliad ar gyfer deallusrwydd artiffisial. Roedd llawer yn defnyddio prosesau a chanllawiau arfer gorau gan eu cyrff sector neu fasnach megis techUK. Roedd rhai yn teimlo bod y niferoedd cynyddol o brosesau asesu risg sefydliadol gwahanol yn adlewyrchu diffyg canllawiau ac arferion gorau cyffredinol gan y llywodraeth. Ymhlith yr ymatebwyr hyn, awgrymodd llawer y byddai'n ddefnyddiol i fusnesau ddeall safbwynt y llywodraeth ar arferion gorau sy'n gysylltiedig â deallusrwydd artiffisial, gyda rhai yn argymhell canllaw canolog ar ddefnyddio deallusrwydd artiffisial yn ddiogel.

79. Nododd llawer o'r ymatebwyr eu bod yn cydymffurfio â'r fframweithiau cyfreithiol presennol sy'n cofnodi risgiau sy'n gysylltiedig â deallusrwydd artiffisial, megis diogelwch cynhyrchion a diogelu data personol. Nododd ymatebwyr y dylai unrhyw fesurau deallusrwydd artiffisial yn y dyfodol osgoi dyblygu neu wrthddweud rheolau a chyfreithiau sy'n bodoli eisoes.

80. Pwysleisiodd ymatebwyr yn gyson pa mor bwysig oedd tryloywder, gyda rhai yn tynnu sylw at adnoddau rhannu gwybodaeth fel cardiau model. Yn debyg i Gwestiwn 2, awgrymodd rhai ymatebwyr y byddai labelu'r defnydd o ddeallusrwydd artiffisial yn fuddiol i ddefnyddwyr, yn enwedig o ran datblygu llythrennedd ynglŷn â chynnwys maleisus o bosibl a gynhyrchir drwy ddeallusrwydd artiffisial, megis ffugiadau dwfn a thwyllwybodaeth. Roedd ychydig o'r ymatebwyr yn dadlau y gall labelu deallusrwydd artiffisial helpu i bennu disgwyliadau ynglŷn â gwasanaeth a dylid ei wneud er mwyn diogelu defnyddwyr. Gan adlewyrchu'r atebion i F1, soniodd ymatebwyr hefyd y dylai gwasanaethau fod yn dryloyw ynglŷn â'r data a ddefnyddir i hyfforddi modelau deallusrwydd artiffisial fel y gall defnyddwyr ddeall sut mae adnoddau a gwasanaethau yn gweithio yn ogystal â'u cyfyngiadau.

81. Roedd yr ymatebion yn dangos bod maint sefydliad yn ddylanwad ar y gallu i asesu risgiau sy'n gysylltiedig â deallusrwydd artiffisial. Er bod sefydliadau mwy o faint yn sôn eu bod yn ymgysylltu â chwsmeriaid a chyflenwyr i lunio a rhannu arferion gorau, roedd rhai busnesau bach yn gofyn am ragor o gymorth i asesu risgiau sy'n gysylltiedig â deallusrwydd artiffisial a rhoi'r egwyddorion deallusrwydd artiffisial ar waith yn effeithiol.

## **Modelau sylfaen a'r fframwaith rheoleiddio**

### **F1. Pa heriau penodol y bydd modelau sylfaen megis modelau iaith mawr neu fodolau ffynhonnell agored yn eu hachosi i reoleiddwyr sy'n ceisio pennu cyfrifoldeb cyfreithiol am ganlyniadau deallusrwydd artiffisial?**

82. Er bod ymatebwyr yn cefnogi'r fframwaith rheoleiddio deallusrwydd artiffisial a nodwyd yn y papur gwyn, roedd llawer yn pryderu y gallai fod angen dull rheoleiddio pwrpasol ar gyfer modelau sylfaen. Yn benodol, nododd ymatebwyr mai prif nodwedd modelau sylfaen yw eu cymhlethdod technegol a phwysleisiwyd eu potensial i ategu llawer o gymwysiadau gwahanol mewn sawl sector. Pwysleisiodd bron chwarter yr ymatebwyr fod modelau sylfaen yn ei gwneud yn anodd i bennu cyfrifoldeb cyfreithiol am ganlyniadau deallusrwydd artiffisial, gan rannu achosion defnydd damcaniaethol lle roedd bai ar y rhai i fyny'r gadwyn ac i lawr y gadwyn. Pwysleisiodd ymatebwyr fod diffyg tryloywder technegol, cadwyni cyflenwi cymhleth, ac anghymesureddau o ran gwybodaeth yn atal esboniadwyedd digonol, atebolrwydd, ac asesiadau risg ar gyfer modelau sylfaen.

83. Roedd llawer o'r ymatebwyr yn pryderu am ansawdd y data a ddefnyddir i hyfforddi modelau sylfaen ac a yw'r data hyfforddi yn briodol ar gyfer holl gymwysiadau modelau i lawr y gadwyn. Nododd ymatebwyr nad oedd yn glir a yw'r data a ddefnyddir i hyfforddi modelau sylfaen yn cydymffurfio â chyfreithiau presennol, megis y rhai ar gyfer diogelu data ac eiddo deallusol. Nododd ymatebwyr fod diffyg diffiniadau a safonau ar gyfer data hyfforddi. Roedd ymatebwyr yn teimlo y gellid gwella'r defnydd o ddata drwy fesurau rhannu gwybodaeth gwell, mesuriadau meincnodi a safonau, a neilltuo'r cyfrifoldeb am bennu a yw data yn briodol i gymhwysiad penodol yn glir i berson penodol.

84. Pwysleisiodd rhai ymatebwyr gymhlethdod cadwyni cyflenwi modelau sylfaen, gan ddadlau bod anghymesureddau o ran gwybodaeth rhwng datblygwyr i fyny'r gadwyn (sydd â goruchwyliaeth dechnegol) a chyflwynwyr i lawr y gadwyn (sydd â goruchwyliaeth dros gymwysiadau) nid yn unig yn gwneud y cyfrifoldebau cyfreithiol am ganlyniadau deallusrwydd artiffisial yn annelwig ond hefyd yn atal gweithgareddau monitro a lliniaru risgiau digonol. Er i rai ymatebwyr nodi bod grym y farchnad yn nwylo datblygwyr modelau sylfaen yn bennaf ac awgrymu mai nhw oedd yn y sefyllfa orau i lliniaru risgiau cysylltiedig, roedd eraill yn dadlau mai cyfyngedig fyddai trosolwg datblygwyr dros y risgiau sy'n gysylltiedig â chymwysiadau penodol i lawr y gadwyn. Cododd llawer bryderon ynglŷn â diffyg mesurau i farnu'n drwyadl i ba raddau roedd model sylfaen yn briodol i gymhwysiad penodol.

85. Nododd ychydig o'r ymatebwyr bryderon ynglŷn â mynediad ehangach at ddeallusrwydd artiffisial, gan gynnwys ffynhonnell agored, rhyddhau gwybodaeth yn amhriodol, neu ddefnydd maleisus. Fodd bynnag, nododd nifer tebyg o ymatebwyr bwysigrwydd ffynhonnell agored i arloesi, tryloywder, ac ymddiriedaeth ym maes deallusrwydd artiffisial.

## **F2. A ydych chi'n cytuno bod mesur lefelau cyfrifiannu yn cynnig adnodd posibl y gellid ei ystyried fel rhan o drefniadau llywodraethu modelau sylfaen?**

86. Roedd hanner yr ymatebwyr yn teimlo bod cyfrifiannu yn ddirprwy annigonol ar gyfer gofynion llywodraethu, gyda llawer yn dadlau y byddai trothwyon sy'n gysylltiedig â chyfrifiannu yn dyddio'n gyflym oherwydd pa mor gyflym y mae technoleg yn newid. Fodd bynnag, roedd bron hanner yn cytuno i ryw raddau y byddai mesur lefelau cyfrifiannu yn ddefnyddiol wrth lywodraethu modelau sylfaen, gan awgrymu y gellid ei ddefnyddio i asesu a ddylai model deallusrwydd artiffisial penodol ddilyn gofynion penodol pan oedd yn cael ei ddefnyddio ar y cyd â mesurau llywodraethu eraill. Nododd ychydig o'r ymatebwyr y byddai mesur lefelau cyfrifiannu yn un ffordd o gofnodi effaith amgylcheddol gwahanol fodolau deallusrwydd artiffisial.

## **F3. A oes unrhyw ddulliau eraill o lywodraethu modelau sylfaen a fyddai'n fwy effeithiol?**

87. Cafwyd cefnogaeth gyffredinol i fesurau ac adnoddau llywodraethu ar gyfer deallusrwydd artiffisial dibynadwy, gyda'r ymatebwyr yn dadlau o blaid defnyddio trefniadau llywodraethu sefydliadol, safonau technegol, a thechnegau sicrwydd sy'n benodol ar gyfer modelau sylfaen.

88. Argymhellodd rhai ymatebwyr y dylid asesu galluoedd a chymwysiadau modelau sylfaen yn hytrach na chyfrifiannu. Roedd ymatebwyr yn teimlo y byddai mesurau dilysu modelau, megis archwiliadau a gwerthusiadau, yn effeithiol, gyda rhai yn awgrymu y dylai'r rhain fod yn ofynion gorfodol. Nododd rhai ymatebwyr bwysigrwydd monitro i lawr y gadwyn neu wylriadwriaeth ar ôl cyrraedd y farchnad. Awgrymodd un ymatebydd bwll tywod cyn cyflwyno.

89. Roedd nifer bach o ymatebwyr am weld gofynion statudol ar fodelau sylfaen. Roedd ychydig yn cefnogi moratoria, gwaharddiadau, neu gyfyngiadau ar fodelau sylfaenol a'u defnyddiau. Awgrymodd eraill y dylid defnyddio contractau, trwyddedau, a chytundebau â defnyddwyr, gyda'r ymatebwyr hefyd yn nodi pwysigrwydd mesurau ffisegol a seiberddiogelwch.

## Pyllau tywod a meiciau arbrofi deallusrwydd artiffisial

### **S1. Pa un o'r modelau pwll tywod a ddisgrifir yn adran 3.3.4 fyddai'n fwyaf tebygol o gefnogi arloesedd?**

90. Er bod mwyafrif mawr o'r ymatebwyr yn gefnogol iawn i byllau tywod yn gyffredinol, ystyriwyd mai modelau "amlysector, amlyreoleiddiwr" (MSMR) "un sector, amlyreoleiddiwr" (SSMR) fyddai'n fwyaf tebygol o gefnogi arloesedd.

91. Roedd dros draean o'r ymatebwyr yn teimlo y byddai'r model MSMR yn cefnogi arloesedd, gan nodi y byddai'r sail draws-sector yn galluogi rheoleiddwyr i ddatblygu canllawiau effeithiol mewn ymateb i faterion byw, cysoni rheolau, cydgysylltu prosesau gweithredu, sicrhau a ellir cymhwyso modelau at sectorau lle mae diogelwch yn hollbwysig, a nodi ysgogwyr polisi ategol. Awgrymodd ymatebwyr y dylai pwll tywod MSMR fynd i'r afael â materion sy'n gysylltiedig â rhoi'r egwyddorion deallusrwydd artiffisial ar waith, gan gynnwys nodi unrhyw fylchau yn y fframwaith, unrhyw orgyffwrdd â'r trefniadau rheoleiddio presennol, unrhyw heriau o ran cydgysylltu rhwng sectorau a rheoleiddwyr, ac unrhyw ffactorau sy'n atal y fframwaith rheoleiddio rhag cael ei roi ar waith yn effeithiol, megis adnoddau rheoleiddwyr, ac ymdrin â nhw. Pwysleisiodd ymatebwyr hefyd y dylai'r pwll tywod fod yn hyblyg ac addasadwy er mwyn diogelu datblygiadau technolegol newydd rhag datblygiadau yn y dyfodol.

92. Roedd yr un nifer o ymatebwyr yn cefnogi'r model SSMR. Nododd ymatebwyr y byddai'n haws lansio'r modelau SSMR a'r MSSR am fod eu prosesau cydgysylltu rhwng un sector neu reoleiddiwr yn symlach. Am y rheswm hwn, roedd ymatebwyr yn teimlo mai'r modelau hyn fyddai sbarduno'r gwerth mwyaf uniongyrchol. Awgrymodd rhai y gellid addasu pwll tywod un sector neu un rheoleiddiwr yn fodel MSMR wrth i waith fynd rhagddo er mwyn sicrhau manteision y ddau fodel.

### **S2. Beth y gallai'r llywodraeth ei wneud i sicrhau manteision mwyaf posibl pyllau tywod i arloeswyr deallusrwydd artiffisial?**

93. Roedd rhai ymatebwyr yn dadlau y dylai'r pwll tywod gael ei ddatblygu a'i gyflwyno ar y cyd â busnesau, rheoleiddwyr, grwpiau defnyddwyr, ac academyddion ac arbenigwyr eraill. Awgrymodd ymatebwyr y dylid adeiladu ar gryfderau presennol tirwedd reoleiddio'r DU, megis hwyluso gwersi traws-sector drwy'r Fforwm Cydweithredu Rheoleiddio Digidol.

94. Nododd ymatebwyr y dylai'r pwll tywod ddatblygu canllawiau, rhannu gwybodaeth ac adnoddau, a rhoi cymorth i arloeswyr deallusrwydd artiffisial. Yn benodol, dywedodd ymatebwyr y dylai gwybodaeth am gyfleoedd i gymryd rhan gael ei rhannu, gan nodi y byddai rhannu canlyniadau yn annog mwy i gyfranogi. Roedd ymatebwyr am i'r pwll tywod fod yn agored ac yn dryloyw, gyda llawer yn dadlau o blaid sicrhau bod prosesau'r pwll tywod, asesiadau ac adroddiadau rheoleiddio, prosesau penderfyniadau, adolygiadau o dystiolaeth, a chanfyddiadau canlyniadol ar gael i'r cyhoedd. Awgrymodd ymatebwyr y byddai adroddiadau a chanllawiau rheolaidd gan y pwll tywod yn llywio arloeswyr a threfniadau rheoleiddio yn y dyfodol drwy greu prosesau "busnes fel arfer". Roedd

ymatebwyr yn teimlo y dylid cymryd camau i sicrhau bod y pwll tywod mor hygyrch â phosibl, gydag ychydig yn dadlau o blaid llwybrau penodedig a hyfforddiant i fusnesau bach.

95. Roedd ymatebwyr yn teimlo y dylai'r pwll tywod gael ei ddefnyddio i lywio a datblygu safonau technegol a thechnegau sicrwydd y gellir eu defnyddio'n eang. Soniodd ychydig y byddai hyn yn helpu i hyrwyddo arferion gorau mewn diwydiant. Nododd eraill y dylai'r pwll tywod gyd-fynd yn dda â threfniadau rheoleiddio ehangach ar gyfer deallusrwydd artiffisial er mwyn sicrhau'r budd mwyaf. Nododd ymatebwyr hefyd fod pwll tywod yn cynnig cyfle i'r DU ddangos arweinyddiaeth fyd-eang ym maes rheoleiddio deallusrwydd artiffisial a safonau technegol drwy rannu canfyddiadau ac arferion gorau yn rhyngwladol.

96. Nododd ymatebwyr y gallai'r pwll tywod gefnogi arloesedd drwy gynnig manteision ar y farchnad, megis ardystio cynhyrchion, er mwyn sicrhau'r manteision mwyaf i arloeswyr deallusrwydd artiffisial. Ymhlith y cymhellion ariannol eraill a awgrymwyd gan ymatebwyr roedd grantiau arloesi, credydau treth, a chyfranogiad am ddim neu gyfranogiad a ariennir mewn pyllau tywod amgylchedd profi o dan oruchwyliaeth. Roedd ychydig o randdeiliaid yn cytuno y byddai cyllid yn helpu egin fusnesau a busnesau bach sydd â llai o adnoddau sefydliadol i gymryd rhan mewn pyllau tywod sy'n canolbwyntio ar ymchwil a datblygu. Awgrymodd ymatebwyr y gallai'r pwll tywod gydweithio â chwmnïau buddsoddi yn y DU ac yn rhyngwladol i greu cyfleoedd i gwmnïau sy'n cymryd rhan.

### **S3. Beth y gallai'r llywodraeth ei wneud i hwyluso cyfranogiad mewn pwll tywod rheoleiddio deallusrwydd artiffisial?**

97. Awgrymodd rhai ymatebwyr y byddai grantiau, cymorthdaliadau, a chredydau treth yn annog busnesau bach ac egin fusnesau i gymryd rhan mewn modelau pwll tywod sy'n defnyddio llawer o adnoddau ac sy'n canolbwyntio ar ymchwil a datblygu, megis amgylcheddau profi o dan oruchwyliaeth.

98. Roedd ymatebwyr yn cefnogi amrywiaeth o gymhellion i hwyluso'r broses o gymryd rhan mewn modelau pwll tywod gwahanol, gan gynnwys mynediad at setiau data safonol ac wedi'u hanonymyddio, a chynlluniau achredu a fyddai'n dangos cysondeb â gofynion rheoleiddio a helpu i sicrhau mynediad at y farchnad. Cafwyd rhywfaint o gefnogaeth i gystadlaethau arloesi a fyddai'n helpu i ddewis cyfranogwyr.

99. Yn debyg i S2, roedd ymatebwyr yn cytuno y byddai cydweithio ac ymgynghori ag amrywiaeth o randdeiliaid yn helpu i hwyluso cyfranogiad eang. Awgrymodd ymatebwyr ganolfannau ymchwil, rhaglenni sbarduno, a phartneriaethau â phrifysgolion. Cafwyd cefnogaeth o blaid cynnwys grŵp amrywiol o randdeiliaid yn ystod camau cynnar datblygu pwll tywod, yn enwedig er mwyn nodi meysydd rheoleiddio â risg uchel. Cafwyd rhywfaint o gefnogaeth i fframweithiau gwerthuso wedi'u cysoni rhwng sectorau er mwyn lleihau'r baich rheoleiddio ac ennyn mwy o ddiddordeb ymhlith darpar randdeiliaid. Cynigiodd un ymatebydd lwyfan ar-lein penodedig a fyddai'n rhoi mynediad at ganllawiau perthnasol a rhoi porth i gyflwyno ac olrhain ceisiadau ynghyd â fforwm cymunedol.

100. Cafwyd cefnogaeth eang i broses ymgeisio syml gyda chanllawiau clir, templedi, a gwybodaeth am gymhwystra a gofynion cyfreithiol. Mynegodd ymatebwyr gefnogaeth i feini prawf ymuno ac ymadael clir, gan nodi pwysigrwydd lleihau'r baich gweinyddol ar fusnesau bach ac egin fusnesau er mwyn lleihau'r rhwystr i gymryd rhan.

#### **S4. Pa sectorau diwydiannol neu ddosbarthiadau o gynnyrch a fyddai'n cael y budd mwyaf o bwll tywod deallusrwydd artiffisial?**

101. Er nad oedd unrhyw gonsensws cyffredinol ynglŷn â sector na dosbarth o gynnyrch penodol a fyddai'n cael y budd mwyaf o bwll tywod deallusrwydd artiffisial, nododd ymatebwyr ddau sector lle roedd diogelwch yn hollbwysig sydd â chryn risg bosibl: gofal iechyd a thrafnidiaeth. Nododd ymatebwyr fod y sectorau hyn wedi'u nodweddu gan anallu i brofi yn y byd go iawn ac y byddai pwll tywod deallusrwydd artiffisial yn fuddiol iddynt. Nododd ymatebwyr y potensial i wella canlyniadau gofal iechyd, diogelwch cleifion, a chydymffurfiaeth â chanllawiau ar breifatrwydd cleifion drwy feithrin arloesi mewn meysydd megis adnoddau diagnostig, meddyginiaethau wedi'u personoli, darganfod cyffuriau, a dyfeisiau meddygol. Nododd ymatebwyr eraill dwf cerbydau awtonomaidd a systemau trafndiaeth deallus eraill ynghyd â brwdfrydedd mawr o du'r diwydiant i brofi'r fframwaith rheoleiddio.

102. Awgrymodd rhai ymatebwyr y byddai gwasanaethau ariannol ac yswiriant yn cael budd o bwll tywod deallusrwydd artiffisial oherwydd buddsoddiad mawr y sector mewn awtomeiddio a deallusrwydd artiffisial. Nododd ymatebwyr hefyd fod gwasanaethau ariannol ac yswiriant yn cael eu goruchwyllo gan sawl rheoleiddiwr, gan gynnwys Swyddfa'r Comisiynydd Gwybodaeth, yr Awdurdod Rheoleiddio Darbodus, yr Awdurdod Ymddygiad Ariannol, a'r Rheoleiddiwr Pensiynau. Nododd ymatebwyr y gallai gwasanaethau ariannol ddefnyddio pwll tywod deallusrwydd artiffisial i archwilio cymwysiadau deallusrwydd artiffisial ar gyfer asesu risg, canfod twyll, masnachu algorithmig, a gwasanaeth cwsmeriaid.

103. Nodwyd gan un ymatebydd fod y sector niwclear eisoes yn cael budd o bwll tywod deallusrwydd artiffisial. Mae'r Swyddfa Rheoleiddio Niwclear ac Asiantaeth yr Amgylchedd wedi dysgu gwersi o'u pwll tywod rheoleiddio eu hunain er mwyn datblygu'r cysyniad o bwll tywod deallusrwydd artiffisial rhyngwladol ar gyfer y sector niwclear.



# Atodiad D: Crynodeb o dystiolaeth asesu effaith

Mae'r atodiad hwn yn rhoi crynodeb o'r dystiolaeth ysgrifenedig a gawsom mewn ymateb i'n hymgyngoriad ar asesu effaith rheoleiddio deallusrwydd artiffisial.<sup>124</sup> Gofynnwyd wyth cwestiwn, gan gynnwys saith cwestiwn agored neu led-agored a ddenodd amrywiaeth o fyfyrddodau ysgrifenedig. Gofynnwyd:

1. A ydych chi'n cytuno bod y rhesymeg dros ymyrryd yn ymdrin yn gynhwysfawr â niweidiau presennol a niweidiau yn y dyfodol ac yn dangos tystiolaeth ohonynt?
2. A ydych chi'n cytuno bod ymddiriediaeth gynyddol yn ffactor sylweddol sy'n ysgogi'r galw am systemau deallusrwydd artiffisial?
3. A oes gennych unrhyw dystiolaeth ychwanegol i gefnogi'r amcangyfrifon a'r tybiaethau canlynol ar gyfer y fframwaith?
4. A ydych chi'n cytuno â'r amcangyfrifon sy'n gysylltiedig â'r swyddogaethau canolog?
5. A ydych chi'n ymwybodol o unrhyw fetrigau amgen i fesur amcanion polisi?
6. A ydych chi'n credu y byddai rhai systemau deallusrwydd artiffisial yn cael eu gwahardd o dan Opsiynau 1 a 2, oherwydd craffu rheoleiddiol cynyddol?
7. A ydych chi'n cytuno â'n hasesiad o bob opsiwn polisi yn erbyn yr amcanion?
8. A oes gennych unrhyw dystiolaeth ychwanegol sy'n profi neu'n gwrthbrofi ein dadansoddiad yn yr asesiad o'r effaith?

Cawsom 64 o ymatebion ysgrifenedig i gyd i'r ymgynghoriad ar yr asesiad o'r effaith gan sefydliadau ac unigolion. Disgrifir ein dull dadansoddi yn Atodiad A a cheir crynodeb o'r ymatebion i'r cwestiynau hyn isod.

## **Cwestiwn 1: A ydych chi'n cytuno bod y rhesymeg dros ymyrryd yn ymdrin yn gynhwysfawr â niweidiau presennol a niweidiau yn y dyfodol ac yn dangos tystiolaeth ohonynt?**

Crynodeb o'r ymatebion:

Roedd dros hanner yr ymatebwyr yn anghytuno bod y rhesymeg dros ymyrryd yn ymdrin yn gynhwysfawr â'r dystiolaeth o niweidiau presennol a niweidiau yn y dyfodol. Nododd bron hanner yr ymatebwyr nad ymdrinnir yn ddigonol â phob risg. Roedd llawer o'r ymatebwyr hyn yn dadlau nad yw'r rhesymeg yn cyfrif am niweidiau annisgwyl na risgiau dirfodol a systemig. Roedd un ymatebydd yn dadlau nad yw'r rhesymeg yn ystyried effaith deallusrwydd artiffisial ar hawliau dynol. Awgrymodd un ymatebydd y dylai fod gofynion gorfodol o ran casglu data'n foeseogol a dadleuodd un arall o blaid mesurau cyn cyflwyno er mwyn lliniaru risgiau deallusrwydd artiffisial.

Awgrymodd dros chwarter yr ymatebwyr y dylid dadansoddi risgiau a chyfleoedd ar gyfer pob sector. Roedd yr ymatebwyr hyn yn aml yn dadlau na chyfrifir am niweidiau na manteision posibl mewn diwydiannau gwahanol, megis effaith deallusrwydd artiffisial ar swyddi.

---

<sup>124</sup> [UK Artificial Intelligence Regulation Impact Assessment](#), yr Adran dros Wyddoniaeth, Arloesedd a Thechnoleg, 2023.

Roedd rhai ymatebwyr yn dadlau y dylai'r llywodraeth adeiladu ar y dystiolaeth ynglŷn â niweidiau presennol a niweidiau yn y dyfodol yn ogystal ag ymyriadau posibl. Pwysleisiodd llawer o'r ymatebwyr hyn bwysigrwydd cynnwys safbwyntiau amrywiol a llais y cyhoedd wrth gynnal ymchwil a rheoleiddio deallusrwydd artiffisial.

Nododd ychydig o'r ymatebwyr y dylai'r llywodraeth a'r rheoleiddwyr fabwysiadu dull hyblyg sy'n monitro ac yn gallu ymaddasu i ddatblygiadau technolegol.

Nododd ychydig o'r ymatebwyr y bydd gormod o reoleiddio ac ymyrryd gan y llywodraeth yn llesteirio arloesedd yn hytrach na'i annog. Roedd yr ymatebwyr hyn yn dadlau bod angen cydbwyo risgiau lliniaru a galluogi manteision deallusrwydd artiffisial.

Nododd un ymatebydd y dylai fod rheoleiddiwr annibynnol dros ddeallusrwydd artiffisial.

## **Cwestiwn 2: A ydych chi'n cytuno bod ymddiriediaeth gynyddol yn ffactor sylweddol sy'n ysgogi'r galw am systemau deallusrwydd artiffisial?**

Crynodeb o'r ymatebion:

Cytunodd dros hanner yr ymatebwyr fod ymddiriedaeth yn ffactor sylweddol sy'n ysgogi'r galw am systemau deallusrwydd artiffisial. Fodd bynnag, roedd tua chwarter yn anghytuno ac roedd rhai yn ansicr.

Rhoddodd dros draean o'r ymatebwyr ateb ysgrifenedig a allai gynnig rhagor o ddirnadaethau y tu allan i gytuno neu anghytuno. O'r rhain, pwysleisiodd llawer o'r ymatebwyr mai tryloywder, addysg a mesurau llywodraethu (megis rheoleiddio a safonau technegol) sy'n cynyddu ymddiriedaeth. Adlewyrchwyd y syniadau hyn ymhlith yr ymatebwyr a oedd yn cytuno ac yn anghytuno bod ymddiriedaeth yn ysgogi'r galw am ddeallusrwydd artiffisial.

Roedd ymatebwyr hefyd yn dadlau y gallai ymddiriedaeth mewn deallusrwydd artiffisial gael ei gwanhau oherwydd pryderon ynglŷn â thuedd neu ddiogelwch. Nododd rhai o'r ymatebwyr hyn fod tuedd annheg neu anhryloyw mewn systemau deallusrwydd artiffisial nid yn unig yn lleihau ymddiriedaeth ond yn effeithio fwyaf ar gymunedau sydd eisoes wedi'u hymyleiddio. Roedd rhai ymatebwyr yn dadlau y byddai blaenoriaeth ar arloesedd ar draul ymddiriedaeth mewn dull rheoleiddio yn lleihau ymddiriedaeth.

O blith yr ymatebwyr a oedd yn anghytuno bod ymddiriedaeth yn ysgogi'r galw am ddeallusrwydd artiffisial ac a oedd wedi cyflwyno ymatebion ysgrifenedig pellach, daeth dwy brif thema i'r amlwg: Yn gyntaf, mae'r galw am ddeallusrwydd artiffisial yn cael ei ysgogi gan gymhellion economaidd ac ariannol ac, yn ail, mae'n cael ei ysgogi gan

ddatblygiadau technolegol. Er enghraifft, nododd un ymatebydd y gallai deallusrwydd artiffisial wella cynhyrchiant ac felly broffidoldeb cwmnïau. Tynnodd ymatebwyr sylw at ddatblygiadau technolegol hefyd fel ffactor sy'n ysgogi'r galw am ddeallusrwydd artiffisial, gyda dau ymatebydd yn nodi y gallai "ofni colli cyfle" mewn technolegau newydd ymhlith cwmnïau ysgogi eu galw am systemau deallusrwydd artiffisial.

Roedd yr ymatebwyr a oedd yn anghytuno yn aml yn awgrymu bod y galw cynyddol am ddeallusrwydd artiffisial a'r cyfraddau mabwysiadu cynyddol yn digwydd ar draul diogelu'r cyhoedd a lliniaru risgiau.

### **Cwestiwn 3: A oes gennych unrhyw dystiolaeth ychwanegol i gefnogi'r amcangyfrifon a'r tybiaethau canlynol ar gyfer y fframwaith?**

Crynodeb o'r ymatebion:

Ymatebwyd i bob datganiad yn wahanol. Cafwyd cytundeb i raddau gwahanol â phob datganiad. Mewn adborth ysgrifenedig, awgrymodd rhai ymatebwyr fod ein hamcangyfrifon a'n tybiaethau yn dibynnu ar ffactorau cymhleth neu nad oes modd llunio amcangyfrifon ynglŷn â deallusrwydd artiffisial am fod gormod o elfennau ansicr.

O ran yr amcangyfrif cyntaf, sef y bydd mabwysiadu/defnyddio deallusrwydd artiffisial yn effeithio ar 431,671 o fusnesau, yn llai na'r 3,170 o fusnesau, yn ôl yr amcangyfrif, sy'n cyflenwi/cynhyrchu deallusrwydd artiffisial, roedd yr ymatebwyr a oedd yn anghytuno o'r farn ei fod yn tanddatgan nifer y busnesau y bydd deallusrwydd artiffisial yn debygol o effeithio arnynt, y gall

y niferoedd newid yn gyflym am ei bod yn hawdd integreiddio deallusrwydd artiffisial mewn cynnyrch neu wasanaeth, bod y rhaniad rhwng y rhai sy'n mabwysiadu deallusrwydd artiffisial a'r rhai sy'n ei gynhyrchu braidd yn artiffisial, ac y dylid ystyried defnyddwyr hefyd.

O ran yr ail ddatganiad, sef y bydd y rhai sy'n mabwysiadu/defnyddio cynhyrchion a gwasanaethau deallusrwydd artiffisial yn wynebu costau is na'r rhai sy'n cynhyrchu a/neu'n cyflenwi cynhyrchion a gwasanaethau sy'n cynnig datrysiadau deallusrwydd artiffisial, cafwyd rhywfaint o anghytuno ac un ymateb yn cytuno. Roedd y rhai a anghytunodd â'r datganiad yn dadlau na fydd defnyddwyr deallusrwydd artiffisial yn wynebu costau is na chynhyrchwyr deallusrwydd artiffisial oherwydd y gall defnyddwyr yn fwy cyffredinol wynebu costau (cynyddol) defnyddio cymwysiadu deallusrwydd artiffisial. Ar y llaw arall, soniodd un ymatebydd y bydd arbedion cost yn gymwys i ddefnyddwyr heb ddealltwriaeth ddwfn o'r dechnoleg ac y bydd cynhyrchwyr yn wynebu costau cyflogau uchel oherwydd cronfa fach o weithwyr sy'n gallu gweithredu systemau deallusrwydd artiffisial uwch.

O ran y trydydd amcangyfrif y byddai costau ymgysfarwyddo (sy'n cyfeirio yma at y gost i fusnesau am uwchsgilio cyflogaion mewn cyfundrefn reoleiddio newydd) o tua £2.7m

i £33.7m, nododd un neu ddau o'r ymatebwyr a anghytunodd y gallai costau ymgysfarwyddo amrywio o fusnes i fusnes. Roedd yr ymatebwyr hyn yn dadlau bod yr amrediad cyfredol yn tanddatgan y costau llawn, gan argymhell y dylid ystyried costau eraill. Awgrymodd rhai fod angen i ddefnyddwyr gael eu hyfforddi ar risg weddilliol a sut i oresgyn tuedd awtomatiaeth. Soniodd eraill y bydd archwiliadau annibynnol o systemau deallusrwydd artiffisial yn creu llawer o swyddi newydd i gyflogaion tra hyfforddedig.

Yn olaf, o ran y pedwerydd amcangyfrif y bydd costau cydymffurfio (sy'n adlewyrchu yma y gost i fusnesau am addasu elfennau o'r busnes i gydymffurfio â safonau newydd) o tua £107 miliwn i £6.7 biliwn, cafwyd ragor o anghytundeb. Dywedodd rhai ymatebwyr y dylai costau cydymffurfio fod mor isel â phosibl, ond nid oedd unrhyw gytundeb ynglŷn â'r ffordd orau o gyflawni hyn. Nododd ymatebwyr eraill na fydd cwmnïau yn cydymffurfio ac y byddai cydymffurfiaeth yn golygu bod angen ymgymryd â gweithgareddau busnes newydd.

### **Cwestiwn 4: A ydych chi'n cytuno â'r amcangyfrifon sy'n gysylltiedig â'r swyddogaethau canolog?**

Crynodeb o'r ymatebion:

Roedd mwyafrif bach iawn o'r ymatebwyr yn anghytuno i ryw raddau â'r amcangyfrifon a amlinellwyd yn yr asesiad effaith rheoleiddio ar gyfer deallusrwydd artiffisial, gan awgrymu bod yr amcangyfrifon ar gyfer y swyddogaeth ganolog yn rhy uchel.

Soniodd rhai ymatebwyr y gallai'r swyddogaeth ganolog gyflwyno deallusrwydd artiffisial a defnyddio awtomatiaeth i sicrhau effeithlonrwydd a lleihau amcangyfrifon costau. Nododd dau ymatebydd hefyd y gallai'r swyddogaeth ganolog ddefnyddio technegau megis dysgu rhwng cymheiriaid a rhwydweithiau i leihau amcangyfrifon costau.

Ar y llaw arall, nododd rhai ymatebwyr fod yr amcangyfrifon ar gyfer y swyddogaeth ganolog yn rhy isel. Roedd rhai ymatebwyr yn credu bod yr amcangyfrifon cyfredol yn rhy isel am nad ydynt yn cyfrif am gostau sy'n gysylltiedig ag uwchsgilio cyflogeion y swyddogaeth ganolog yn hwyr. Awgrymodd un ymatebydd y byddai'r galw cynyddol am ddeallusrwydd artiffisial ymhlith y sector masnachol yn cynyddu costau ymhellach, ac yn creu heriau yn y swyddogaeth ganolog o ran cael gafael ar ddatrysiadau deallusrwydd artiffisial oherwydd pwysau cost chwyddiannol. Awgrymodd rhai ymatebwyr, oherwydd graddfa a galluoedd cynyddol deallusrwydd artiffisial, y byddai angen swyddogaeth ganolog fwy o faint i reoleiddio'r dechnoleg, gan ddadlau bod y costau cyfredol yn debygol o fod yn rhai ceidwadol.

Cytunodd ychydig o'r ymatebwyr fod yr amcangyfrifon yn gywir. Fodd bynnag, nododd llawer y byddai'n heriol pennu ffigur penodol ar gyfer yr amcangyfrifon sy'n gysylltiedig â'r swyddogaeth ganolog, gan awgrymu bod diffyg eglurder wrth ddiffinio termau yn ei gwneud yn anodd asesu cywirdeb yr amcangyfrifon.

### **Cwestiwn 5: A ydych chi'n ymwybodol o unrhyw fetrigau amgen i fesur amcanion polisi?**

Crynodeb o'r ymatebion:

Awgrymodd dros draean o'r ymatebwyr fetrigau amgen y gellid eu defnyddio i fesur yr amcanion polisi. Roedd rhai awgrymiadau yn cynnwys olrhain nifer y modelau sy'n cael eu harchwilio o ran tuedd a thegwch; nifer y digwyddiadau ynglŷn â deallusrwydd artiffisial y rhoddir gwybod amdanynt ac yr ymchwilir iddynt; a metrigau sy'n gysylltiedig â'r ffordd y mae'r fframwaith yn gweithredu megis nifer y rheoleiddwyr sy'n cyhoeddi canllawiau, natur y canllawiau a'r canlyniadau cysylltiedig i sefydliadau sydd wedi eu mabwysiadu, neu ddangosyddion sentiment i randdeiliaid. Ymhlith yr awgrymiadau eraill roedd olrhain ymddiriedaeth y cyhoedd a'i barodrwydd i dderbyn systemau deallusrwydd artiffisial.

Awgrymodd tua chwarter yr ymatebwyr fframweithiau a modelau sy'n bodoli eisoes. Awgrymodd un neu ddau o'r ymatebwyr y byddai trefniadau effeithiol i asesu a rheoleiddio niwed yn allweddol i fesur yr amcanion polisi.

### **Cwestiwn 6: A ydych chi'n credu y byddai rhai systemau deallusrwydd artiffisial yn cael eu gwahardd o dan Opsiynau 1 a 2, oherwydd craffu rheoleiddiol cynyddol?**

Crynodeb o'r ymatebion:

Roedd dros hanner yr ymatebwyr yn cytuno y byddai rhai systemau deallusrwydd artiffisial yn cael eu gwahardd o dan Opsiynau 1 a 2 oherwydd craffu rheoleiddiol cynyddol. Roedd tua chwarter yr ymatebwyr yn anghytuno ac roedd ychydig o dan draean yn ansicr.

O blith yr ymatebwyr a ymhelaethodd ar eu syniadau, awgrymodd traean fod rhai systemau deallusrwydd artiffisial yn fygythiad i gymdeithas ac y dylid eu gwahardd. Pwysleisiodd yr ymatebwyr hyn y byddai gwaharddiad yn lleihau risgiau deallusrwydd artiffisial, ac yn cael effaith gadarnhaol, yn eu barn nhw. Awgrymodd rhai y byddai'r fframwaith rheoleiddio yn methu pe na bai'n gosod gwaharddiadau.

Awgrymodd rhai rhanddeiliaid y byddai rhai systemau deallusrwydd artiffisial yn cael eu gwahardd. Fodd bynnag, awgrymodd nifer tebyg na fyddai'r trefniadau craffu rheoleiddiol o dan Opsiynau 1 a 2 yn ddigon i wahardd systemau deallusrwydd artiffisial. Roedd y ddwy

set hyn o ymatebion yn adlewyrchu dealltwriaeth groes ynghylch pa mor ddwys oedd y rheoliadau arfaethedig, yn hytrach na safbwyntiau cynhenid ynglŷn â sut y gallai rheoliadau effeithio ar y sector. Nododd ychydig nad oedd yr asesiad o'r effaith yn gallu rhoi digon o dystiolaeth ynglŷn â pha systemau deallusrwydd artiffisial a allai gael eu gwahardd.

### **Cwestiwn 7: A ydych chi'n cytuno â'n hasesiad o bob opsiwn polisi yn erbyn yr amcanion?**

Crynodeb o'r ymatebion:

Roedd ychydig dros draean o'r ymatebwyr naill ai'n cytuno'n gryf neu'n cytuno i ryw raddau â'r asesiad o bob opsiwn polisi yn erbyn amcanion, gyda'r rhan fwyaf yn ymateb eu bod yn cytuno i ryw raddau. Roedd nifer tebyg naill ai'n anghytuno'n gryf neu'n anghytuno i ryw raddau, gyda'r rhan fwyaf o'r rhain yn ymateb dim ond i ryw raddau roeddent yn anghytuno. Nid oedd tua chwarter o'r ymatebwyr yn cytuno nac yn anghytuno, neu roeddent wedi nodi eu bod yn ansicr.

### **Cwestiwn 8: A oes gennych unrhyw dystiolaeth ychwanegol sy'n profi neu'n gwrthbrofi ein dadansoddiad yn yr asesiad o'r effaith?**

Crynodeb o'r ymatebion:

Awgrymodd tua hanner yr ymatebion ysgrifenedig nad oedd yr asesiad effaith rheoleiddio ar gyfer deallusrwydd artiffisial yn amcangyfrif effeithiau deallusrwydd artiffisial yn ddigonol. Nododd yr ymatebwyr hyn fod effeithiau deallusrwydd artiffisial yn llawer mwy sylweddol a mwy niweidiol na'r hyn a awgrymwyd gan yr asesiad effaith rheoleiddio ar gyfer deallusrwydd artiffisial a'r papur gwyn.

Nododd ychydig o dan draean y dylai'r llywodraeth weithredu'n ddi-oed i reoleiddio technolegau deallusrwydd artiffisial datblygol. Pwysleisiodd yr ymatebwyr hyn y dylai gweithredu'n amserol fod yn ffocws allweddol rheoleiddio deallusrwydd artiffisial, o ystyried pa mor gyflym y mae galloedd y dechnoleg yn datblygu.

Nododd rhai ymatebwyr fod gormod o ansicrwydd i wneud asesiadau cywir. Roedd yr ymatebwyr hyn yn meddwl y byddai unrhyw amcangyfrif yn anghywir oherwydd natur deallusrwydd artiffisial a'r ansicrwydd mawr ynglŷn â datblygiadau yn y dyfodol.

Awgrymodd rhai ymatebwyr y dylai'r rheoleiddwyr gysoni'r ffordd y maent yn ymdrin â deallusrwydd artiffisial, gan bwysleisio bod y defnydd o'r technolegau hyn rhwng sectorau yn gofyn am reoleiddio cydgysylltiedig a chyson.

