

Rhif Ymchwil Gymdeithasol: 30/2025

Dyddiad cyhoeddi: 02/04/2025

Moeseg mewn ymchwil data gweinyddol: adolygiad llenyddiaeth

Mae'r ddogfen yma hefyd ar gael yn Saesneg.

OGL © Hawlfraint y Goron ISBN Digidol 978-1-83715-528-6

Moeseg mewn ymchwil data gweinyddol: adolygiad llenyddiaeth

Awdur: Emma Yates

Adroddiad Ymchwil Llawn: Yates, E. 2025. Moeseg mewn ymchwil data gweinyddol: adolygiad llenyddiaeth. Caerdydd: Llywodraeth Cymru, rhif adroddiad GSR 30/2025.
Ar gael yn: <https://www.llyw.cymru/moeseg-mewn-ymchwil-data-gweinyddol>

Barn yr ymchwilwyr a fynegir yn yr adroddiad ac nid o reidrwydd barn Llywodraeth Cymru.

I gael rhagor o wybodaeth cysylltwch â:

Enw: Emma Yates

Caffael a Chysylltu Data ar gyfer Ymchwil

Gwasanaethau Gwybodaeth a Dadansoddi

Llywodraeth Cymru

Parc Cathays

Caerdydd

CF10 3NQ

E-bost: ADRWales@llyw.cymru

Tabl cynnwys

Geirfa	2
1. Crynodeb Gweithredol.....	6
2. Cyflwyniad	9
3. Nodau.....	11
4. Methodoleg	12
5. Canfyddiadau	14
Llenyddiaeth Gefndir	14
Mapio Fframweithiau Moesegol, Egwyddorion a Chanllawiau	23
Llywodraethu Da	26
Budd y Cyhoedd.....	32
Ymddiriedaeth	36
Ymgysylltu â'r Cyhoedd.....	38
6. Casgliadau	41
7. Argymhellion o'r Llenyddiaeth	43
Cyfeiriadau.....	45
Atodiad A	50
Atodiad B	54
Atodiad C	57

Geirfa

Atebolrwydd

Y 'mecanweithiau llywodraethu a goruchwylio effeithiol ar gyfer unrhyw brosiect' (Swyddfa Ddigidol a Data Ganolog 2020).

Data gweinyddol

Data a gesglir pan fydd pobl yn defnyddio gwasanaethau cyhoeddus, megis ym myd addysg, y gwasanaeth iechyd, neu'r llysoedd neu system fudd-daliadau. Mae'r data hyn yn cael eu coladu gan y llywodraeth i ganiatáu ar gyfer darparu gwasanaethau ac nid ydynt wedi'u hanelu'n bennaf at ddibenion ymchwil.

YDG

Ymchwil Data Gweinyddol.

ADR DU

Mae Ymchwil Data Gweinyddol y DU yn cynnwys pedair partneriaeth genedlaethol (ADR England, ADR Northern Ireland, ADR Scotland ac YDG Cymru), a'r Swyddfa Ystadegau Gwladol. Mae'n gweithredu i sicrhau bod ymchwilwyr yn cael mynediad at ddata Llywodraeth y DU ar ffurf ddiogel er mwyn creu mewnwelediadau am gymdeithas i lywio'r gwaith o lunio polisiau. Ariennir ADR UK gan y Cyngor Ymchwil Economaidd a Chymdeithasol, fel rhan o Ymchwil ac Arloesi'r DU.

ADR Cymru

Mae Ymchwil Data Gweinyddol Cymru yn uno arbenigwyr ym mhrifysgolion Abertawe a Chaerdydd, Banc Data SAIL, ac ystadegwyr, economegwyr ac ymchwilwyr cymdeithasol o Lywodraeth Cymru. Nod y data a gysylltir ac a ddadansoddir gan ADR Cymru yw darparu tystiolaeth ar gyfer y meysydd blaenoriaeth yng Nghymru.

ALF

Maes Cyswllt Dienw

Anhysbys

Y broses o dynnu dynodwyr allweddol (fel enwau a dyddiadau geni) o ddata personol.

Cyswllt Data

Dod â data ynghyd o ddwy ffynhonnell wahanol neu fwy sy'n ymwneud â'r un unigolyn, teulu, lle neu ddigwyddiad.

Rhannu data

Symud gwybodaeth neu ddata o un sefydliad neu adran i un arall.

DEA

Deddf yr Economi Ddigidol 2017.

Dat-adnabod data

Mae data'n cael eu dat-adnabod trwy ddileu gwybodaeth bersonol a fyddai'n caniatáu i unigolion gael eu hadnabod yn uniongyrchol.

Fframwaith pum egwyddor ddiogel

Set o egwyddorion arfer gorau a ddefnyddir i lywodraethu mynediad diogel at ddata ar gyfer ymchwil. Y pum egwyddor ddiogel yw: data diogel (risg datgeliad wedi'i reoli), prosiectau diogel (defnyddir data'n briodol ac er lles y cyhoedd), pobl ddiogel (mae ymchwilwyr wedi'u hyfforddi'n briodol), gosodiadau diogel (mae defnydd anawdurdodedig wedi'i gyfyngu gan y cyfleuster mynediad), ac allbynnau diogel (mae allbynnau'n cael eu sgrinio ar gyfer risg datgelu).

GDPR

Rheoliad Diogelu Data Cyffredinol.

GSR

Ymchwil Gymdeithasol y Llywodraeth.

GSS

Gwasanaeth Ystadegol y Llywodraeth.

ICO

Swyddfa'r Comisiynydd Gwybodaeth.

Llywodraethu gwybodaeth

Fframwaith neu gyfres o weithdrefnau i reoli'r ffordd yr ymdrinnir â gwybodaeth er mwyn sicrhau bod prosesau data yn ddiogel ac yn cydymffurfio â'r holl rwymedigaethau moesegol a chyfreithiol.

NSDEC

Pwyllgor Ymgynghorol Moeseg Data yr Ystadegydd Gwladol.

OECD

Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd.

SYG

Swyddfa Ystadegau Gwladol.

OSR

Swyddfa Rheoleiddio Ystadegau.

Budd cyhoeddus

Yn aml yn cael ei ddefnyddio'n gyfnewidiol â 'budd y cyhoedd' a 'lles y cyhoedd', mae budd cyhoeddus yn ymwneud â manteision ymchwil a dadansoddi i'r cyhoedd.

Ymgysylltu â'r cyhoedd set bwrpasol o weithgareddau wedi'u cynllunio i hyrwyddo deialog barhaus gyda'r cyhoedd am ymchwil data gweinyddol, wedi'i sbarduno gan wrando gweithredol ac ymateb (gwefan ADR UK).

SAIL

Banc Data Cyswllt Diogel Gwybodaeth Ddienw wedi'i leoli ym Mhrifysgol Abertawe. Mae'n cynnwys ystod o setiau data heb eu nodi y gellir eu cysylltu ar gyfer prosiectau ymchwil cymeradwy gan ymchwilwyr achrededig.

Panel Defnyddwyr SAIL

Panel cyhoeddus sy'n cynnwys tua 16 aelod o'r cyhoedd sy'n byw ledled Cymru ac yn dod at ei gilydd i gynghori ymchwilwyr ar eu prosiectau SAIL.

Atebolrwydd cymdeithasol

Atebolrwydd yn deillio o ymgysylltu â phobl mewn cymdeithas.

Trwydded gymdeithasol

Dealltwriaeth neu set o ddisgwyliadau rhwng y wladwriaeth a dinasyddion ynghylch sut y defnyddir data sy'n mynd y tu hwnt i reoleiddio ffurfiol.

Tryloywder

Camau gweithredu, prosesau a data yn cael eu gwneud yn agored i'w harchwilio drwy gyhoeddi gwybodaeth am y prosiect mewn fformat cyflawn, agored, dealladwy, hygyrch, a rhad ac am ddim (Swyddfa Ddigidol a Data Ganolog 2020).

TRE

Amgylcheddau Ymchwil Dibynadwy (amgylcheddau cyfrifiadurol hynod ddiogel sy'n darparu mynediad o bell at ddata i ymchwilwyr cymeradwy eu defnyddio mewn ymchwil).

UKSA

Awdurdod Ystadegau'r Deyrnas Unedig.

1. Crynodeb gweithredol

Cyflwyniad

- 1.1 Cynhaliwyd yr adolygiad hwn o lenyddiaeth i lywio gwaith ymchwilwyr y llywodraeth ac academyddion ar draws partneriaeth Ymchwil Data Gweinyddol (YDG) Cymru. Mae'n asesu'r llenyddiaeth foeseg ddiweddar sy'n berthnasol i ymchwil data gweinyddol i nodi ac archwilio themâu allweddol.
- 1.2 Mae YDG Cymru yn dod â gwyddonwyr data, academyddion a thimau arbenigol yn Llywodraeth Cymru ynghyd i hwyluso a chynnal ymchwil gan ddefnyddio data gweinyddol a chynhyrchu tystiolaeth i lywio polisiau a phenderfyniadau yng Nghymru. Mae YDG Cymru yn rhan o bartneriaeth ADR UK fwy. Nod y bartneriaeth yw gwneud y mwyaf o ddefnyddioldeb data dat-adnabyddedig i helpu i lunio'r gwasanaethau cyhoeddus a ddarperir ac yn y pen draw ei nod yw gwella bywydau pobl yng Nghymru ac yn ehangach.

Cefndir

- 1.3 Mae'n anodd diffinio moeseg data yn fanwl gywir. Mae'n ystyried elfennau cymdeithasol, technegol a gweithdrefnol mentrau data ac ymchwil, gan gynnwys asesu arferion data i ddeall eu heffeithiau cadarnhaol a negyddol (Open Data Institute 2025). Mae hefyd yn cynnwys cysyniadau sylfaenol fel atebolrwydd, tegwch, tryloywder, budd y cyhoedd, ymgysylltu â'r cyhoedd ac ymddiriedaeth (OECD 2019). Mae'n broses weithredol sy'n berthnasol ar bob cam o'r broses ymchwil ac sy'n benodol i'r cyd-destun (Y Gymdeithas Frenhinol 2020).
- 1.4 O fewn ymchwil data gweinyddol mae rhai pryderon moesegol penodol yn cynnwys: terfynau dat-adnabod data (Laurie a Stevens 2016) ac ymwybyddiaeth isel y cyhoedd o gysylltu data ac ymchwil data gweinyddol (OSR 2023). Mewn data gweinyddol, nid yw'r cysylltiad rhwng unigolyn yn rhannu data drwy ddefnydd gwasanaeth cyhoeddus, defnydd ymlaen a sefydliadau sy'n defnyddio'r data ar gyfer ymchwil bob amser yn amlwg (ONS 2023). Roedd Sexton et al. 2017 yn disgrifio data'n cael eu 'hestyn i ffwrdd o'u cyd-destun gwreiddiol' (Sexton et al. 2017).
- 1.5 Bellach mae gan ymchwilwyr fynediad at lefelau digyngsail o ddata ar lefel poblogaeth, a gesglir at ddibenion gweinyddol, y gellir eu cysylltu a'u dadansoddi i greu mewnwelediadau pellgyrhaeddol am sut mae cymdeithas yn

gweithredu ar wahanol lefelau. Mae ymchwil data gweinyddol yn rhan o fframwaith cymdeithasol a thechnegol cyfoes o ddigideiddio cyflym. Mae natur seilwaith data mawr yn ei gwneud yn anodd iawn i unigolion gadw arolygiaeth dros ansawdd, effaith gymdeithasol neu fudd cyhoeddus ymchwil (Leonelli 2020). Mae anghydbwysedd pŵer yn yr ecosystem ddata ehangach gydag unigolion â rheolaeth gyfyngedig neu wybodaeth am ddefnydd eilaidd o'u data (Ada Lovelace Institute 2022).

- 1.6 Mae yna lu o egwyddorion moeseg, canllawiau a dogfennaeth a gall hyn fod yn ddryslyd ac yn anodd ei lywio (Kalkman et al. 2019). Mae fframweithiau, offer a chanllawiau moesegol yn cynnwys asesiadau moesegol arddull 'rhestr wirio' yn ogystal ag egwyddorion moesegol cyfannol mwy cyffredinol.

Atebolrwydd

- 1.7 Mae angen i atebolrwydd ymgorffori atebolrwydd cymdeithasol fel nad yw wedi'i gyfyngu i brosesau goruchwyllo ffurfiol (Cyngor Biofoeseg Nuffield 2015). Mae'r llenyddiaeth yn amlinellu bod angen esboniad o'r hyn sy'n digwydd i ddata pobl a pham (Cyngor Biofoeseg Nuffield 2015), er mwyn i bobl allu arfer lefel o reolaeth. Mae angen i atebolrwydd hefyd gynnwys proses ar gyfer deall niwed posibl yn ogystal â manteision defnyddio data, gwneud iawn i unigolion pan aiff pethau o chwith, a gweithredu systemau ar gyfer adrodd am ganggymeriadau o fewn sefydliadau (Kennedy et al. 2020).

Tryloywder

- 1.8 Mae llawer o fanteision i gynyddu tryloywder mewn ymchwil data gweinyddol, ond mae heriau a risgiau'n gysylltiedig â hyn, megis darparu'r lefel gywir o wybodaeth glir ar bynciau cymhleth ar gyfer gwahanol gynulleidfaoedd (Britain Thinks 2021).

Budd y Cyhoedd

- 1.9 Mae budd cyhoeddus, neu fantais wedi'i diffinio'n glir yn ystyriaeth foesegol allweddol wrth gychwyn ar unrhyw ymchwil. Fodd bynnag, mae 'budd cyhoeddus' yn derm goddrychol a hyblyg a rhaid ei ddiffinio gan awdurdodau cyhoeddus yn ôl cyd-destun (Laurie a Stevens 2016). Mae angen cyfathrebu budd y cyhoedd mewn prosiectau ymchwil yn eang a bod yn gynhwysol mewn trafodaethau am yr hyn y mae budd y cyhoedd yn ei olygu (Ada Lovelace Institute 2022 a Kennedy et al. 2022).

Ymgysylltu â'r Cyhoedd

- 1.10 Cydnabyddir yn gyffredinol bod angen mwy o ymgysylltu â'r cyhoedd ar rannu data ac ymchwil cysylltu data (OSR ac ADR UK 2022 a Harkness et al. 2022). Mae gwahanol lefelau o gyfranogiad i'w hystyried er mwyn sicrhau bod y cyhoedd yn cael eu cynrychioli, yn gallu cymryd rhan yn ystyrion mewn deialog am ddefnyddio data a chymryd rhan mewn penderfyniadau am ddata ac ymchwil er budd y cyhoedd. Gellid defnyddio dulliau ymchwil ansoddol i gynorthwyo gydag ymgysylltiad cyhoeddus (Aitkin et al. 2016).

Trwydded Gymdeithasol

- 1.11 Mae'r llenyddiaeth yn cyfeirio at drwydded gymdeithasol ar gyfer defnyddio data fel bod set o ddisgwyliadau cyhoeddus a ddeallir yn eang ynghylch sut y bydd y llywodraeth yn defnyddio data ar gyfer ymchwil, gan nad yw hyn yn glir ar hyn o bryd (Canolfan Moeseg Data ac Arloesi 2020). Mae'n gysyniad a ddefnyddir i egluro dealltwriaeth rhwng y wladwriaeth a dinasyddion ynghylch sut mae adnoddau cyhoeddus yn cael eu defnyddio. Er enghraifft, mae dealltwriaeth gyffredinol ynghylch sut y caiff cyllid cyhoeddus ei wario. Gall trwydded gymdeithasol gynnwys gweithgareddau sy'n mynd y tu hwnt i gydymffurfio â deddfwriaeth a gallai fod gwrthwynebiad neu her gymdeithasol os na chaiff y disgwyliadau hynny eu bodloni (Carter et al. 2015).

Cyfyngiadau

- 1.12 Mae hwn yn adolygiad llenyddiaeth thematig dethol sy'n archwiliadol gyda'r nod o ddistyllu themâu allweddol yn hytrach nag adolygu'n systematig yr holl lenyddiaeth ar y pwnc hwn. Oherwydd hynny, mae'n darparu sylfaen mewn pynciau moeseg allweddol yn hytrach nag adolygiad arbenigol neu fanwl.

2. Cyflwyniad

- 2.1 Mae moeseg ymchwil yn rhan bwysig a chymhleth o ddefnyddio data ar gyfer ymchwil. Ar gyfer ymchwilwyr y llywodraeth ac academyddion mae llawer o egwyddorion, fframweithiau, codau ymarfer a chanllawiau i'w hystyried a'u cymhwyso yn ystod y broses ymchwil, yn ogystal â gofynion penodol i'r sefydliad. Gall maint yr arweiniad sydd ar gael fod yn llethol a dryslyd wrth i sefydliadau, partneriaethau a phroffesiynau ddatblygu eu canllawiau a'u hoffer moesegol eu hunain fwyfwy. Fel yr amlygir isod, cydnabyddir yn gyffredinol yn y llenyddiaeth bod yna lawer o sefydliadau sy'n cyhoeddi canllawiau ar gynnal ymchwil i gymuned ymchwil y DU.
- 2.2 Mae cangen o ymchwil sy'n dadansoddi ac yn cysylltu setiau data gweinyddol mawr o wahanol rannau o'r llywodraeth. Mae'r term 'data gweinyddol' yn cyfeirio at wybodaeth a gesglir pan fydd pobl yn defnyddio gwasanaethau cyhoeddus, megis mewn addysg, y gwasanaeth iechyd, neu'r llysoedd neu'r system budd-daliadau. Mae'r data hyn yn cael eu coladu gan y llywodraeth i ganiatáu ar gyfer darparu gwasanaethau ac nid ydynt wedi'u hanelu'n bennaf at ddibenion ymchwil. Mae rhannu data ag Amgylcheddau Ymchwil Dibynadwy (TREs) yn galluogi ymchwilwyr i ddadansoddi a chysylltu setiau data o wahanol adrannau. Mae'r Banc Data Cyswllt Diogel Gwybodaeth Ddienw (SAIL) yn enghraifft o TRE a chadwrfa ddata sy'n dal llawer o setiau data gweinyddol wedi'u dat-adnabod y gallir eu cysylltu ar gyfer ymchwil er budd y cyhoedd. Mae data gweinyddol yn adnodd cyhoeddus sy'n helpu i ddarparu gwasanaethau cyhoeddus ond hefyd gallir eu defnyddio i greu dealltwriaeth a all helpu gyda phenderfyniadau'r llywodraeth a datblygu polisi. Mae data gweinyddol yn werthfawr iawn at ddibenion ymchwil oherwydd y potensial hwn i gynhyrchu mewnwelediadau ar lefel y boblogaeth pan gânt eu dadansoddi a'u cysylltu â setiau data eraill o wahanol rannau o'r llywodraeth. Fodd bynnag, mae defnyddio data gweinyddol at ddibenion ymchwil hefyd yn codi llawer o gwestiynau moesegol a chyfreithiol.
- 2.3 Gall fod yn anodd dirnad y canllawiau moesegol perthnasol i ymgynghori â hwy yn ystod ymchwil data mawr a sut i'w cymhwyso'n ymarferol. Mae moeseg data, fel cangen o foeseg ymchwil, yn faes sy'n tyfu, yn enwedig o ystyried datblygiad a defnydd cynyddol o dechnoleg AI. Mae Sefydliad Ada Lovelace yn nodi:

Yn y llenyddiaeth, mae lluo o egwyddorion, fframweithiau a chanllawiau ar foeseg AI wedi dechrau cydgyfeirio o amgylch egwyddorion fel tryloywder, cyfiawnder, tegwch, rhadlonrwydd, cyfrifoldeb a phreifatrwydd. Fodd bynnag, nid oes ymdrech fyd-eang wedi'i gwneud eto i drosi'r egwyddorion hyn yn arferion moeseg ymchwil AI, nac i benderfynu sut y dylai sefydliadau ymchwil ddehongli neu weithredu egwyddorion moesegol. Mae hyn yn ei gwneud yn ofynnol i ymchwilwyr ystyried dehongliadau a dealltwriaeth foesegol amrywiol mewn rhanbarthau (Ada Lovelace Institute 2022).

- 2.4 Mae hyn hefyd wedi cael ei gydnabod wrth rannu data mewn ymchwil ieuchyd, fe wnaeth Kalkman et al. (2019) nodi 'digonedd o egwyddorion a normau' ond 'gydag amrywiad sylweddol yn y geiriad a lefel y manylder, nifer a chynnwys y normau yr ystyrir eu bod yn angenrheidiol i ddiogelu egwyddor, a'r dulliau cyd-destunol o ddefnyddio egwyddorion a normau' (Kalkman et al. 2019). Maent yn crynhoi bod y cwestiwn yn dod yn sut i gysoni'r egwyddorion hyn yn fframwaith llywodraethu a moesegol sy'n hwyluso rhannu data ac ymddiriedaeth y cyhoedd.
- 2.5 O ystyried arwyddocâd cynyddol moeseg data a'r lluo o fframweithiau moeseg sy'n bodoli, mae'r adolygiad hwn o llenyddiaeth yn asesu'r llenyddiaeth foeseg ddiweddar sy'n berthnasol i ymchwil data gweinyddol i nodi ac archwilio themâu sy'n berthnasol i ddadansoddwyr ac ymchwilwyr y llywodraeth sy'n defnyddio data'r llywodraeth.
- 2.6 Nodir nodau a methodoleg yr adolygiad hwn yn Adrannau 3 a 4 isod. Dilynnir y rhain gan drafodaeth ar y llenyddiaeth gefndir yn Adran 5, gan gynnwys diffiniadau perthnasol o foeseg data, gwybodaeth gyd-destunol am ddigideiddio mewn cymdeithas, a sut mae data gweinyddol wedi dod yn rhan o'r dirwedd ymchwil. Mae'r rhan fwyaf o ganfyddiadau'r adolygiad yn parhau yn adran 5, a drefnir o dan y penawdau a ganlyn: Mapio Fframweithiau Moesegol, Llywodraethu Da, Budd y Cyhoedd, Ymddiriedaeth, ac Ymgysylltu â'r Cyhoedd. Yn olaf, mae rhai casgliadau ac argymhellion cryno yn adrannau 6 a 7.

3. Nodau

3.1 Nodau'r adolygiad hwn yw:

- archwilio'r sylfaen dystiolaeth ddiweddar yn y llenyddiaeth ar foeseg ymchwil data gweinyddol
- nodi fframweithiau ac egwyddorion moesegol perthnasol; a
- nodi themâu a gwneud argymhellion ar arfer gorau yn seiliedig ar y dystiolaeth.

4. Methodoleg

- 4.1 Roedd yr adolygiad llenyddiaeth yn seiliedig ar set gychwynnol o chwiliadau llenyddiaeth a gynhaliwyd ar-lein rhwng 25/01/2023 a 17/02/2023, gan ddefnyddio nifer o gronfeydd data academiaidd a gwefannau dethol perthnasol. Cwblhawyd chwiliadau cychwynnol gan Wasanaeth Llyfrgelloedd Llywodraeth Cymru. Defnyddiwyd geiriau allweddol gyda ffocws ar foeseg data, uniondeb, 'lles y cyhoedd' a chyswllt data ac ymchwil data gweinyddol. Cyflwynir tabl manwl o'r strategaethau chwilio, gan gynnwys ffynonellau a thermau chwilio, yn Atodiad A.
- 4.2 Ategwyd y chwiliad hwn gan lenyddiaeth arall a nodwyd o chwiliadau ar y gwefannau canlynol: Sefydliad Ada Lovelace, y Gymdeithas Frenhinol, ADR UK, Awdurdod Ystadegau'r DU (UKSA) a gwefannau'r llywodraeth. Ategwyd hyn ymhellach trwy fynychu seminarau, gweminarau a digwyddiadau ar-lein perthnasol sy'n berthnasol i foeseg data mewn ymchwil data gweinyddol yn 2024, gan gynnwys seminar Civil Service Live ar foeseg data a gweminar ar foeseg a diogelu'r cod ymarfer ar gyfer ystadegau a ddarperir gan y Swyddfa Rheoleiddio Ystadegau (OSR) at y dyfodol.
- 4.3 Cyfyngwyd y canlyniadau i bapurau ymchwil diweddar (2010 ymlaen), papurau ymchwil y llywodraeth, erthyglau cyfnodolion a pheth llenyddiaeth 'lwyd'. At ddiben yr adolygiad hwn, mae llenyddiaeth lwyd yn golygu 'lenyddiaeth a gynhyrchir gan sefydliadau annibynnol, cymdeithas sifil, y trydydd sector, sefydliadau llywodraethol neu fasnachol neu gan academyddion ar gyfer cynulleidfaoedd anacademaidd' (Kennedy et al. 2020). Dylid ystyried y wybodaeth a gyflwynir yn y ddogfen hon yng nghyd-destun y meini prawf chwilio a ddefnyddiwyd.
- 4.4 Nodwyd 29 o ddogfennau gan y chwiliad llenyddiaeth llyfrgell (Atodiad B), ac ategwyd hyn gan 29 o ddogfennau pellach o chwiliadau atodol (fel y rhestrir yn Atodiad C). Ni adolygwyd 8 dogfen oherwydd nad oedd modd eu cyrchu, cawsant eu cyhoeddi mewn cyfnodolion llai perthnasol, nid oeddent yn canolbwyntio ar Gymru na'r DU, neu wedi'u cyhoeddi cyn 2010. Yn y pen draw, dewiswyd 50 o ddogfennau ar gyfer yr adolygiad hwn. Er na ddefnyddiwyd system bwysoli ffurfiol ar gyfer y llenyddiaeth a nodwyd, rhoddwyd mwy o bwyslais yn gyffredinol ar ddogfennau mwy diweddar lle mae'r prif ffocws yn benodol ar foeseg data a lle mae'r cyd-destun yn ymchwil data gweinyddol.

Cyfyngiadau

- 4.5 Adolygiad thematig o lenyddiaeth yw hwn ac felly nid yw'n gynhwysfawr. Cydnabyddir bod llawer iawn o lenyddiaeth moeseg o wahanol sectorau a sefydliadau. Mae hwn yn adolygiad llenyddiaeth dethol sy'n canolbwyntio ar y llenyddiaeth sydd fwyaf perthnasol i ymchwilwyr a dadansoddwyr y llywodraeth.
- 4.6 Fe'i cynhaliwyd yn ystod 2023 a 2024 i ganiatáu ar gyfer myfyrio ac ar y cyd ag astudiaeth beilot gyda Phanel Defnyddwyr SAIL i archwilio moeseg data trwy ddau weithdy trafod (Yates et al. 2025 – i'w gyhoeddi tua diwedd Ebrill 2025). Dewiswyd themâu ar sail eu mynychder yn y llenyddiaeth, ond proses iteraidd a dethol oedd hon yn hytrach nag un systematig, ac ysgrifennir y canfyddiadau yn unol â'r dull hwn. Roedd y ffocws ar archwilio cysyniadau sylfaenol allweddol sydd fwyaf perthnasol i ymchwil data gweinyddol o safbwynt y llywodraeth mewn ffordd gyfannol ac mae llawer o hyblygrwydd deongliadol i lawer o themâu. Ffocws yr adolygiad oedd deall materion moesegol sy'n dod i'r amlwg ac mae llawer o'r rhain yn ymwneud â chwestiynau mwy a ffactorau cyd-destunol ynghylch ymddiriedaeth yn y llywodraeth, digideiddio ac effaith gymdeithasol datblygiadau technoleg cyflym, ac yn cael eu llunio ganddynt.
- 4.7 Nid yw'r holl ddogfennau a nodwyd yn y chwiliadau yn cael eu trafod yn yr adran Canfyddiadau isod. Fodd bynnag, gellir gweld yr holl ddogfennau a nodwyd yn y chwiliadau yn y tabl canlyniadau yn Atodiadau B ac C.

5. Canfyddiadau

- 5.1 Mae'r adran hon yn dechrau gydag archwiliad o'r llenyddiaeth gefndir sy'n nodi diffiniadau a gwybodaeth gyd-destunol am ddata mewn cymdeithas a'i effaith a'i berthnasedd i ymchwil data gweinyddol. Mae'n parhau ag adolygiad o'r llenyddiaeth sy'n manylu ar y materion moesegol penodol sy'n nodwedd o ymchwil data gweinyddol, megis ymwybyddiaeth gyhoeddus isel. Mae'r canfyddiadau'n parhau gyda thrafodaeth fer ar fframweithiau, egwyddorion ac offer moeseg data sydd fwyaf perthnasol i ymchwilwyr y llywodraeth. Yn olaf, trafodir y prif themâu a nodir o'r llenyddiaeth fel a ganlyn: llywodraethu da (tegwch, atebolrwydd a thryloywder), budd y cyhoedd, ymddiriedaeth ac ymgysylltu â'r cyhoedd.

Llenyddiaeth gefndir

Beth yw moeseg data?

- 5.2 Mae moeseg data a'r egwyddorion a'r prosesau sy'n ymwneud â phob cam o'r daith ymchwil (gan gynnwys sut mae data'n cael eu casglu, eu prosesu, eu curadu, eu storio, eu rhannu, eu defnyddio a'u dileu) yn rhan gynyddol bwysig o foeseg ymchwil wrth i ddigideiddio ddod yn fwy eang. Mae'n gangen o foeseg ymchwil sy'n esblygu gyda datblygiadau ystadegol a thechnolegol a chanfyddiad cymdeithasol a chyhoeddus ynghylch sut y dylid defnyddio data ar gyfer ymchwil:

Mae'r parth moeseg data yn edrych ar agweddau ac effeithiau moesegol a chymdeithasol casglu data, prosesu data, dadansoddi data a chymhwyso'r dadansoddiad hwnnw. Gallai ystyriaethau moesegol felly amrywio o archwilio cysyniadau sylfaenol megis caniatâd, tegwch ac atebolrwydd, ac archwilio'r dulliau dadansoddol eu hunain (fel algorithmau penodol), i archwilio effeithiau ehangach systemau a yrrir gan ddata (Y Gymdeithas Frenhinol 2020).

- 5.3 Mae moeseg data hefyd wedi'i diffinio gan yr academyddion Floridi a Taddeo yn eu herthygl: 'Beth yw Moeseg Data?' fel:

...cangen newydd o foeseg sy'n astudio ac yn gwerthuso problemau moesol sy'n ymwneud â data (gan gynnwys cynhyrchu, cofnodi, curadu, prosesu, lledaenu, rhannu a defnyddio), algorithmau (gan gynnwys deallusrwydd artifisial, asiantau artifisial, dysgu peiriannau a robotiaid) ac arferion cyfatebol (gan gynnwys arloesi cyfrifol, rhaglennu, hacio a chodau proffesiynol), er

mwyn llunio a chefnogi datrysiadau moesol dda (e.e. ymddygiad cywir neu werthoedd cywir) (Floridi a Tadde 2016).

- 5.4 Mae'r Sefydliad Data Agored yn diffinio moeseg data fel 'cangen o foeseg sy'n gwerthuso arferion data sydd â'r potensial i effeithio'n andwyol ar bobl a chymdeithas – wrth gasglu, rhannu a defnyddio data' (gwefan y Sefydliad Data Agored 2025). Eglurir hyn ymhellach drwy amlinellu bod y ffocws ar werthuso arferion data er mwyn deall yr effaith gadarnhaol a negyddol fel y gallwn gynnal ymarfer cydbwysu i helpu i bennu camau gweithredu yn y dyfodol.
- 5.5 Mae'r diffiniadau hyn yn dangos ei bod yn anodd diffinio moeseg data mewn termau pendant a hollgynhwysol, ond ei fod yn mynd y tu hwnt i'r gofynion cyfreithiol sy'n gysylltiedig â phrosesau ymchwil a rhannu data ac yn ystyried effeithiau cymdeithasol ar bob cam o'r broses ymchwil. Deellir moeseg data yn y cyd-destun ymchwil fel peidio â thrin data mewn unrhyw ffordd a all niweidio unrhyw un yn uniongyrchol neu'n anuniongyrchol, hyd yn oed os yw rhannu neu ddosbarthu data yn gyfreithlon. Mae'n defnyddio egwyddorion sylfaenol fel 'tegwch', 'tryloywder' a 'phreifatrwydd' i lywio'r broses o wneud penderfyniadau. Mae'r Sefydliad ar gyfer Cydweithrediad a Datblygiad Economaidd (OECD) yn fframio moeseg data fel 'cyfannol' ac yn gysylltiedig â'r angen i gynnal ymddiriedaeth y cyhoedd (OECD 2019).

Pwy sy'n Gweithio ym maes Moeseg Data?

- 5.6 Mae'r Gymdeithas Frenhinol yn ei thirwedd llywodraethu data (2020) yn rhoi trosolwg o sefydliadau allweddol y DU sy'n gweithio ym maes llywodraethu data. Maent yn cynnwys sefydliadau cymdeithas sifil a sefydliadau ymchwil a datblygu megis Sefydliad Ada Lovelace, y Sefydliad Data Agored a Sefydliad Alan Turing. Maent hefyd yn cynnwys cyrff llywodraeth ganolog a'r sector cyhoeddus fel y Ganolfan Moeseg Data ac Arloesi (sydd bellach yn rhan o'r Adran Gwyddoniaeth, Arloesi a Thechnoleg) a Swyddfa'r Comisiynydd Gwybodaeth (ICO) (rheoleiddiwr). Mae'r adroddiad tirwedd hefyd yn rhoi cyflwyniad byr i sut mae gwahanol sefydliadau'n mynd i'r afael â gwahanol agweddau ar lywodraethu data, gan gynnwys: moeseg data, preifatrwydd data ac anonymeiddio, rhannu data a diogelu data.
- 5.7 Yng Nghymru, mae datblygiadau digidol y sector cyhoeddus yn ystod y blynyddoedd diwethaf wedi cynnwys lansio'r Ganolfan Gwasanaethau

Cyhoeddus Digidol a phenodi Prif Swyddogion Digidol Llywodraeth Leol a Llywodraeth Cymru. Mae Strategaeth Ddigidol Cymru (2021) yn nodi 6 chenhadaeth sydd 'o'u cymryd gyda'i gilydd, â'r nod o gyflymu buddion arloesi digidol i bobl, gwasanaethau cyhoeddus ac ar draws ein cymuned fusnes' (Strategaeth Ddigidol i Gymru 2021). Mae Safonau Gwasanaeth Digidol Cymru a ddatblygwyd gan y Ganolfan Gwasanaethau Cyhoeddus Digidol yn nodi'r safonau y disgwylir iddynt gael eu bodloni gan wasanaethau digidol newydd neu wedi'u hailgynllunio sy'n cael eu hariannu'n gyhoeddus. Maent yn cynnwys y gofyniad i ystyried moeseg, preifatrwydd a diogelwch drwyddi draw.

Data yn y Gymdeithas

- 5.8 Er mwyn deall y trafodaethau yn y llenyddiaeth ynghylch themâu canolog moeseg data, yn gyntaf mae angen gwerthfawrogi'r newid sylfaenol yn y defnydd o ddata dros y blynyddoedd diwethaf a'r hyn y mae hyn wedi'i olygu i lywodraeth a chymdeithas. Yn ei hanfod, mae datblygiad technolegau digidol wedi newid yn sylfaenol sut mae pobl yn ymgysylltu â gwybodaeth, ac mae hyn yn effeithio ar bobl mewn ffyrdd newydd, gwahanol ac annisgwyl (Y Gymdeithas Frenhinol 2020). Mae'r dirwedd esblygol hon a'r datblygiadau technolegol cyflym yn cael sgil-ffeithiau ar y llywodraeth sy'n ysgogi datblygiadau ym mholisi'r llywodraeth, llywodraethu gwybodaeth, rheoli data a strategaethau ymgysylltu â'r cyhoedd (Y Gymdeithas Frenhinol 2020).
- 5.9 Ar draws sectorau fe fu ymgyrch bwerus tuag at ddata mawr, digideiddio ac awtomeiddio i gynyddu effeithlonrwydd a lleihau costau. O fewn ymchwil data gweinyddol, bu 'ffrwydrad byd-eang yn y defnydd o ddata poblogaeth ar lefel unigol ar gyfer ymchwil gyda datblygiadau technolegol a chynnydd mewn seilweithiau rhannu data' (Jones a Ford 2018). Bellach mae gan ymchwilwyr fynediad at lefelau digysail o ddata lefel poblogaeth y gallir eu cysylltu a'u dadansoddi i greu mewnwelediadau pellgyrhaeddol am sut mae cymdeithas yn gweithredu ar wahanol lefelau. Mae wedi cael ei gydnabod yn eang bod gan hyn botensial enfawr i ddyfnhau ein dealltwriaeth o gymdeithas a mynd i'r afael â heriau cymdeithasol a pholisi. Fodd bynnag, i ba raddau y gall y cyhoedd neu unigolion gael llais yn y broses neu werthuso neu graffu ar fudd cyhoeddus gwaith o'r fath? Mae Leonelli wedi mynd i'r afael â'r cwestiwn hwn:

Mae sianeli cynhyrchu a lledaenu data fel cyfryngau cymdeithasol, cronfeydd data llywodraethol a storfeydd ymchwil yn gweithredu mewn rhwydwaith byd-

eang, rhyng-gysylltiedig a gwasgaredig, y mae ei weithrediad yn gofyn am amrywiaeth eang o sgiliau ac arbenigedd. Mae natur wasgaredig y penderfyniadau sy'n gysylltiedig â datblygu seilweithiau data mawr a dadansoddeg yn ei gwneud yn amhosibl i unrhyw un unigolyn gadw arolygiaeth dros ansawdd, arwyddocâd gwyddonol ac effaith gymdeithasol bosibl y wybodaeth a gynhyrchir (Leonelli 2020).

- 5.10 Ochr yn ochr â'r datblygiadau hyn, fe fu sgandalau data a thoriadau hysbys iawn sydd wedi cael effaith negyddol ar ganfyddiadau o ddefnydd data, nid yn unig o fewn y maes lle maent wedi digwydd, ond ledled yr ecosystem ddata, oherwydd efallai na fydd y cyhoedd yn gwahaniaethu rhwng gwahanol ddibenion defnyddio data, a'r gwahaniaethau rhwng data a ddefnyddir ar gyfer ymchwil er budd y cyhoedd a'r hyn a ddefnyddir yn y sector preifat er budd masnachol. Fe wnaeth Kennedy et al. (2022) yn eu hymchwil ansoddol ganfod bod yna ddryswch a phryder ynghylch cyfranogiad masnachol mewn defnyddiau data'r sector cyhoeddus a diffyg eglurder cyffredinol ynghylch sut mae'r sector masnachol yn cymryd rhan.

Grym a Chyfranogiad

- 5.11 Sefydlwyd Sefydliad annibynnol Ada Lovelace yn 2018 a'i nod yw sicrhau bod data'n gweithio 'i bobl a chymdeithas' (Ada Lovelace Institute 2025). Mae'r Sefydliad yn feirniadol o'r ecosystem ddigidol ehangach, gan ei ddynodi'n 'gamfanteisiol', 'yn ddi-ddychymyg', ac yn 'ddirymusol', oherwydd ffurfiau anghyfartal o rym a diflaenoriaethu hawliau ac asiantaethau unigol. (Sefydliad Ada Lovelace 2022). Mae'r sefydliad yn dadlau bod rhaniad yn bodoli rhwng y rhai sy'n datblygu naratifau o amgylch data (mabwysiadu dull 'o'r brig i lawr'), a gwrthrychau data. Mae'n argymhell y dylid ailystyried prosesau data a'u gwneud yn fwy cydweithredol i wella cyfranogiad y cyhoedd ac annog mwy o gydraddoldeb a chynhwysiant (Ada Lovelace Institute 2021).
- 5.12 Yn eu hadroddiad, Stiwardiaeth Data Cyfranogol, mae Sefydliad Ada Lovelace yn amlinellu 'fframwaith wedi'i seilio ar degwch, cydraddoldeb a chynhwysiant' (Sefydliad Ada Lovelace 2021). Dylai stiwardiaid data fel 'unigolion a sefydliadau sy'n prosesu a defnyddio data ar ran buddiolwyr', ystyried hawliau gwrthrychau'r data, a'u galluogi i gymryd perchnogaeth o'u data eu hunain trwy gael gwybod am brosesau casglu a dadansoddi, a thrwy gymryd rhan fwy gweithredol mewn llywodraethu data (Ada Lovelace Institute 2021 a 2022). Yn

benodol, mae'r Sefydliad yn eirioli dros ymgysylltu â'r cyhoedd ym mhob cam o brosesu data (Ada Lovelace Institute 2021). I grynhoi, mae'r Sefydliad yn argymhell y dylai stiwardiaid data ddadadeiladu'r sbectrwm grym presennol ac ystyried sut y gallai'r cyhoedd chwarae rhan fwy gweithredol, gan gynnwys cael eu hysbysu, ymgynghori â nhw, eu cynnwys, gwrandio arnynt ac yn y pen draw eu grymuso i wneud penderfyniadau. Dylai'r broses hon fod yn wirioneddol gyfranogol a chynnwys y cyhoedd yn ddilys mewn prosesu data i sicrhau bod pob grŵp o fewn cymdeithas, yn enwedig grwpiau sydd wedi'u hymyleiddio a'u tangynrychioli yn nodweddiadol, yn cael eu cynnwys ac yn cael hawliau data (Ada Lovelace Institute 2021).

Data Gweinyddol yn y Gymdeithas

- 5.13 Mae data gweinyddol a gesglir gan y llywodraeth wedi dod yn rhan o'r dirwedd ymchwil data ac o'i ddadansoddi a'i gysylltu â setiau data eraill gall gyfleu stori gyfoethog am gymdeithas ar addysg, iechyd, lles, incwm, cyflogaeth, mudo, materion cyfiawnder cymdeithasol a throseddau. Mae gan ymchwil data gweinyddol y potensial i ddyfnhau ein dealltwriaeth o gymdeithas er mwyn mynd i'r afael yn well â heriau economaidd-gymdeithasol.
- 5.14 Mae partneriaethau rhwng y byd academiaidd, y llywodraeth a chyllidwyr yn ceisio gwireddu potensial data gweinyddol. Mae YDG Cymru yn gydweithrediad o academyddion ym mhrifysgolion Abertawe a Chaerdydd, swyddogion yn Llywodraeth Cymru a staff Banc Data SAIL sydd wedi'u lleoli ym Mhrifysgol Abertawe. Mae YDG Cymru yn gweithio mewn partneriaeth i gyrchu, dadansoddi a chynhyrchu mewnwelediadau o ymchwil sy'n seiliedig ar ddata i gynorthwyo â llunio polisiau a darparu tystiolaeth i gefnogi Llywodraeth Cymru. Mae SAIL yn blatfform eYmchwil Diogel ac yn TRE achrededig o dan Ddeddf yr Economi Ddigidol 2017 lle cedwir setiau data gweinyddol mawr wedi'u dat-adnabod a gellir eu darparu a'u cyfyngu i ymchwilwyr achrededig i weithio ar brosiectau penodol a chymeradwy heb iddynt adael rhwydwaith SAIL.
- 5.15 Ar wahân i agweddau technegol a llywodraethu ymchwil data gweinyddol, mae ffocws cryf ar ansawdd, uniondeb a gwella safonau moesegol fel yr adlewyrchir yn un o'r nodau strategol yn Rhaglen waith YDG Cymru (2022-2026) sy'n anelu at gynnal ansawdd a chywirdeb gwaith YDG Cymru wrth symud ymlaen:

Cydnabuwyd ein gwaith yn ystod cyfnod peilot ADR UK ar gyfer cyflawni ymchwil effeithiol sydd o safon uchel, yn dderbyniol i'r cyhoedd ac yn cael ei weithredu gyda'r parch mwyaf at safonau deddfwriaethol a moesegol. We will look for opportunities to improve this even further (YDG Cymru 2022).

- 5.16 Mae Panel Defnyddwyr SAIL yn un ffordd o asesu derbynioldeb cyhoeddus prosiectau ymchwil. Fe'i sefydlwyd yn 2011 i 'ddarparu llais cyhoeddus a mesur derbynioldeb cymdeithasol' (gwefan SAIL 2024) o'r gwaith a wnaed ym Manc Data SAIL Prifysgol Abertawe i sicrhau bod data heb eu nodi ar gael i ymchwilwyr achrededig. Mae'r Panel yn cynnwys 16 aelod o'r cyhoedd sy'n cyfarfod unwaith y chwarter i gynghori ymchwilwyr a rhoi persbectif cyhoeddus ar wahanol agweddau ar brosiectau ymchwil. Er enghraifft, drwy adolygu gwybodaeth a gynlluniwyd ar gyfer pobl heb unrhyw gefndir mewn ymchwil data. Mae Panel SAIL hefyd yn eiriolwyr ar gyfer ymchwil cysylltu data.
- 5.17 Mae ADR UK a'r OSR wedi ceisio deall canfyddiadau'r cyhoedd o'r hyn y mae 'lles y cyhoedd' yn ei olygu yng nghyd-destun defnyddio data ar gyfer ymchwil trwy ymarfer deialog cyhoeddus. Mae'n nodi bod 'posibiliadau cynyddol i gynhyrchu mewnwelediadau o ddata gweinyddol a gasglwyd gan y cyhoedd, gan wasanaethau cyhoeddus, gan ddod â mwy fyth o angen i gynnwys y cyhoedd mewn sgysiau am ei ddefnydd yn y dyfodol' (OSR ac ADR UK 2022). Mae'r canfyddiadau'n dangos bod y cyhoedd am fod yn rhan o'r ffordd y gwneir penderfyniadau am ddefnyddio data ar gyfer ymchwil. Yn ogystal, roedd y cyfranogwyr o'r farn y dylid lleihau niwed ac awgrymwyd 'ymgynghori ag aelodau o'r cyhoedd, yn enwedig y rheini â phrofiad bywyd, ynghylch y defnydd posibl o ddata ar gyfer ymchwil neu ddehongli patrymau ystadegol' (OSR ac ADR UK 2022).
- 5.18 Un maes na chafodd ei archwilio yn yr ymarfer hwn oedd moeseg data gweinyddol, a sut mae'n esblygu. Er bod llawer o ffocws wedi'i roi yn hanesyddol ar sicrhau bod data'n cael eu defnyddio'n foesebol mewn prosiectau cysylltu data trwy gyfeirio at brosesau, technoleg a deddfwriaeth ddiogel (er enghraifft, cysylltu gan ddefnyddio meysydd dienw, defnyddio'r fframwaith pum coffr a sefydlu'r pŵer cyfreithiol i rannu data), cymharol ychydig o ymchwil a wnaed i farn y cyhoedd am foesebol data mewn ymchwil, na sut i ddatblygu ymddiriedaeth y cyhoedd ac arferion data cyfranogol. Mae hyn wedi'i gydnabod yn y llenyddiaeth. Er enghraifft, mae'r Ganolfan Moeseg ac Arloesi

Data yn nodi mai mynd i'r afael â rhwystrau technegol a chyfreithiol i rannu data fu'r ffocws ond 'cymharol gyfyngedig fu ymdrech gan y llywodraeth a'r sector cyhoeddus ehangach i fynd i'r afael yn benodol ag ymddiriedaeth y cyhoedd' (Canolfan Moeseg Data ac Arloesi 2020). Roedd Muller et al. (2021) yn honni mewn ymchwil iechyd, nad yw cydymffurfio â'r gyfraith yn ddigon i sicrhau ymddiriedaeth y cyhoedd. Yn yr un modd, mae Jones a Ford (2018) yn tynnu sylw at y ffaith bod defnyddio data lefel unigol ar gyfer ymchwil yn gofyn am lawer mwy na chydymffurfio â deddfwriaeth (fel y gofynion o dan ddeddfau diogelu data a bodloni gofynion o dan y pŵer ymchwil yn Neddf yr Economi Ddigidol 2017) a bod y ffactorau sy'n bwydo i mewn i 'drwydded gymdeithasol' ar gyfer gwaith o'r fath yn gymhleth ac yn ddynamig (Jones a Ford 2018). Felly mae derbyniad y cyhoedd ac ymddiriedaeth y cyhoedd yn gydrannau allweddol o ddata gweinyddol a ddefnyddir ar gyfer ymchwil ac archwilir y themâu hyn yn fanylach isod.

- 5.19 Archwiliwyd moeseg data yn benodol gyda Phanel Defnyddwyr SAIL mewn prosiect peilot a gynhaliwyd gan ymchwilwyr Llywodraeth Cymru. Mae canfyddiadau'r gweithdai a gynhaliwyd gyda'r Panel am eu canfyddiad o foeseg data mewn ymchwil data gweinyddol i'w gweld yn yr adroddiad (cyhoeddiad sydd i'w gyhoeddi ddiwedd Ebrill 2025) sy'n dilyn yr un hwn (Yates et al. 2025).

Moeseg Data Gweinyddol

- 5.20 Mae'r ystyriaethau moeseg data yng nghyd-destun data gweinyddol wedi esblygu gyda datblygiadau ystadegol a thechnolegol, yn ogystal â chanfyddiadau cymdeithasol ynghylch sut y dylid defnyddio data ar gyfer ymchwil. I ddechrau, roedd y rhain yn canolbwyntio ar sicrhau bod mesurau a pholisïau diogelwch priodol ar waith. Mae meddwl academiaidd ynghylch moeseg wedi datblygu i nodi egwyddorion lefel uchel a'r gwrthdaro posibl rhyngddynt, megis y gwrthdaro posibl rhwng sicrhau budd cyhoeddus o rannu a chysylltu data ar gyfer ymchwil a diogelu preifatrwydd (Sexton et al. 2017).

Diogelwch, Ail-adnabod a Risgiau

- 5.21 Roedd Stiles et al. (2011) yn eu hadolygiad o argymhellion arfer gorau ar gyfer defnyddio data gweinyddol yn foesegol mewn ymchwil, yn amlygu'r angen am bolisïau diogelwch priodol, gan gynnwys diogelwch technoleg, gweithdrefnau diogelu, a gweithdrefnau ar gyfer hyfforddiant gorfodol, gan arwain at

ymchwilyr yn meddu ar y sgiliau priodol i ddadansoddi'r data a'r hyfforddiant i sicrhau eu bod yn ymdrin â'r ymchwil mewn modd moesegol. Yn ogystal, maent yn amlygu'r angen am broses caniatâd i sicrhau defnydd priodol o ddata ar gyfer ymchwil (Stiles et al. 2011).

- 5.22 Mewn perthynas â chysylltu data yn benodol, mae nifer o faterion moesegol eraill sydd wedi'u nodi mewn llenyddiaeth ddiweddar. Mae cysylltu data ar gyfer ymchwil yn golygu cysylltu gwahanol setiau data â'i gilydd a sicrhau bod data cysylltiedig wedi'u dat-adnabod ar gael i ymchwilyr. Fodd bynnag, ni ellir diystyru ail-adnabod yn llwyr gan nad oes dull 100% diogel o ddiennw. Gellir rhoi gwahanol ddarnau o wybodaeth at ei gilydd a'u croeswrio. Mae'n bosibl na fydd dibyniaeth ar ddat-adnabod data fel ffordd o gyfiawnhau cysylltu data yn bodloni disgwyliadau'r cyhoedd ynghylch eu preifatrwydd yn y cyd-destun hwn neu gallai roi ymdeimlad ffug o ddiogelwch. Mae'r llenyddiaeth yn amlinellu'r her foesegol hon ar gyfer cysylltu data:

Gall cysylltu data sensitif gynyddu'r tebygolrwydd o ail-adnabod. Gall y mewnwelediadau a geir o gysylltu fod â goblygiadau i unigolion nad oeddent yn ymwybodol ohonynt pan rannwyd data gyda chorff yn y sector cyhoeddus yn wreiddiol. Mae angen asesu'r risg o ail-adnabod a'r posibilrwydd o dorri preifatrwydd yn erbyn budd ehangach y cyhoedd (Canolfan Moeseg Data ac Arloesi 2020).

- 5.23 Mae Laurie a Stevens (2016) yn datgan bod ail-adnabod bob amser yn risg yn y dyfodol, a bod yn rhaid adolygu statws data bob amser. Mae awduron wedi cydnabod efallai nad yw'r cyhoedd yn deall yr hyn y mae anonymeiddio yn ei olygu a: 'tra bod astudiaethau wedi nodi agweddau'r cyhoedd tuag at anonymeiddio, nid yw bob amser yn glir beth mae aelodau'r cyhoedd yn ei ddeall yw ystyr anonymeiddio, na'r hyn y maent yn ei weld sydd ei angen.' (Aitkin et al. 2016). Argymhellir felly bod pobl yn cael eu hysbysu am derfynau anhysbysrwydd a chyfrinachedd (Kalkman et al. 2019).

- 5.24 Materion pellach sy'n benodol i ymchwil cysylltu data yw'r potensial ar gyfer gogwydd trwy wallau cysylltu (efallai nad ydynt yn cael eu dosbarthu ar hap ac a allai felly arwain at ragfarn wrth ddadansoddi), gyda'r potensial cysylltiedig ar gyfer effeithiau negyddol ehangach. Gall hyn fod yn arbennig o berthnasol i boblogaethau anodd eu cyrraedd (Sexton et al. 2017).

- 5.25 Yn ogystal, mae gwiliadwriaeth ar lefel cymdeithas wedi'i nodi gan Aitkin et al. (2016) fel un camddefnydd posibl o ddata:

...mae pryderon hefyd am y potensial am effeithiau negyddol ehangach o ganlyniadau ymchwil. Mae'r rhain yn ymwneud ag: y potensial ar gyfer rhannu data neu gysylltu data i alluogi, neu barhau gwiliadwriaeth dorfol a "Cymdeithas y Brawd Mawr" ganfyddedig; y potensial i unigolion neu grwpiau o fewn cymdeithas gael eu labelu o ganlyniad i ymchwil cysylltu data ac i labelu o'r fath arwain at stigma neu driniaeth wahaniaethol, ac i; y potensial i ymchwil sy'n seiliedig ar ddadansoddi setiau data mawr gael ei defnyddio i lywio polisïau neu arferion sydd wedi'u cynllunio "ar gyfer y llw" yn hytrach nag adlewyrchu amgylchiadau ac anghenion unigol (Aitkin et al. 2016).

Ansicrwydd ac Ymwybyddiaeth Isel y Cyhoedd

- 5.26 Mae parodrwydd o fewn rhai cylchoedd yn y sector cyhoeddus i rannu a chysylltu data gweinyddol, ac i'r buddion gael eu gwireddu, ond ar yr un pryd mae diffyg eglurder hanesyddol ynghylch cyfreithlondeb rhannu data ac ansicrwydd ynghylch darlun moeseg sy'n datblygu. Mae rhan o hyn yn cynnwys y diffyg ymwybyddiaeth ac ansicrwydd ynghylch agwedd y cyhoedd at rannu data a chysylltiadau a'u hyder ynddo (OSR 2023). Mae SYG yn nodi: 'Nid yw'r cysylltiad rhwng unigolyn sy'n rhannu data a'r sefydliad sy'n defnyddio'r data bob amser yn uniongyrchol ac yn amlwg, yn enwedig ar gyfer data gweinyddol' (ONS 2023).
- 5.27 Mae peth ymchwil yn awgrymu, er bod ymwybyddiaeth y cyhoedd yn isel, bod cefnogaeth gyhoeddus i ymchwil data gweinyddol er budd y cyhoedd, ond bod amodau yn gysylltiedig â chefnogaeth o'r fath. Mae pryderon ynghylch cyfrinachedd, rheolaeth unigolion dros eu data, defnydd a chamddefnydd o ddata a niwed posibl (ADR UK 2020). Adolygiad systematig o lenyddiaeth academiaidd a gynhaliwyd gan Kalkman et al. (2019) yn amlinellu pedair thema lefel uchel, gyfanredol ar gyfer rhannu data cyfrifol a moesegol fel a ganlyn: buddion a gwerth cymdeithasol, dosbarthiad risgiau, buddion a beichiau, parch at unigolion a grwpiau, ac ymddiriedaeth ac ymgysylltiad y cyhoedd.
- 5.28 I grynhoi, mae heriau moesegol penodol yn gysylltiedig ag ymchwil data gweinyddol. Mae'r rhain yn cynnwys y risg o ail-adnabod, y potensial am

effeithiau negyddol ehangach ymchwil ar lefel y boblogaeth, lefelau isel o ymwybyddiaeth gyhoeddus o gysyniadau allweddol fel gwneud yn anonymeiddio a'r potensial ar gyfer gwrthdaro rhwng egwyddorion allweddol. Mae elfen 'cyhoeddus' moeseg data gweinyddol yn cynnwys ymddiriedaeth y cyhoedd, ymgysylltu â'r cyhoedd a buddion cyhoeddus ymchwil; mae'r rhain yn themâu allweddol ac yn cael eu harchwilio'n fanwl isod. Yn gyntaf fodd bynnag, mae'r canfyddiadau'n parhau gydag adolygiad byr o fframweithiau moesegol.

Mapio fframweithiau moesegol, egwyddorion a chanllawiau

- 5.29 Mae yna fframweithiau moeseg lluosog ac amrywiol, offer a chodau ymarfer i ymchwilwyr a dadansoddwyr y llywodraeth eu hystyried wrth weithio gyda data. Wrth asesu sut y gellir defnyddio'r adnoddau hyn, mae Egwyddorion Arfer Da yr OECD ar gyfer Moeseg Data yn y Sector Cyhoeddus (2019) yn darparu rhywfaint o ganllawiau ar ddiben fframweithiau moesegol ac yn cynghori 'nad yw fframweithiau moesegol data yn disodli, ond yn hytrach yn ategu, cymorth, ac yn rhyng-gysylltiedig ag offerynnau cyfraith caled perthnasol' (OECD 2019). At hynny, mae eu llwyddiant yn mynnu eu bod yn cael eu hymgorffori ym mhrosesau gwneud penderfyniadau'r sector cyhoeddus ac y dylent gael eu mynegi mewn gweithredoedd a phrosesau ar lefel lai a thechnegol, er enghraifft mewn rheolau rheoli data neu ansawdd data (OECD 2019).
- 5.30 Mae Llywodraeth y DU wedi coladu llawer o egwyddorion moeseg data ac AI yn y Dirwedd Canllawiau Moeseg Data ac AI (Yr Adran Gwyddoniaeth, Arloesedd a Thechnoleg a'r Adran dros Ddigidol, Diwylliant, y Cyfryngau a Chwaraeon 2020). Nod y ddogfen hon yw rhoi eglurder i weision sifil sy'n gweithio gyda data a/neu AI. Mae'r ddogfen yn crynhoi 13 o ddogfennau canllaw, gan nodi pwy greodd bob un ohonynt ac yn rhoi disgrifiad byr, y gynulleidfa arfaethedig a dyddiad y fersiwn ddiweddaraf. Nid yw pob un o'r dogfennau a restrir yn berthnasol i'r adolygiad hwn. Mae rhai perthnasol yn cynnwys: y Fframwaith Moeseg Data (Swyddfa Ddigidol a Data Ganolog 2020), offeryn hunanasesu Awdurdod Ystadegau'r DU (UKSA) (UKSA 2022), y Magenta Book (Trysorlys EF 2020) a'r Aqua Book (Trysorlys EM 2015). Trafodir y rhain yn fyr isod. Heb eu cynnwys yn y Dirwedd Canllawiau Moeseg Data ac AI, ond a allai fod yn berthnasol i weision sifil sy'n llywio'r ystyriaethau moesegol ynghylch ymchwil data gweinyddol, mae offer moesegol penodol i'r proffesiwn megis egwyddorion moesegol yr UKSA (UKSA 2022) a'r canllawiau a gyhoeddwyd ar gyfer y rheini

ym mhroffesiwn Ymchwil Gymdeithasol y Llywodraeth (GSR) (Ethical Assurance for Social and Behavioral Research in Government - GSR 2021).

- 5.31 Y tu hwnt i lenyddiaeth y llywodraeth mae llawer o godau ymarfer, offer a chanllawiau eraill. Er enghraifft, mae Swyddfa Ymchwil ac Uniondeb y DU (UKRIO) wedi cyhoeddi eu Cod Ymarfer ar gyfer Ymchwil (2023) a 'gynlluniwyd i annog ymddygiad da' a 'hyrwyddo uniondeb safonau moesegol uchel mewn ymchwil' (UKRIO 2023). Mae Cynfas Moesegol y Sefydliad Data Agored (2021) yn offeryn ar gyfer unrhyw un sy'n gweithio gyda data i helpu i nodi a rheoli materion moesegol trwy gydol cylch bywyd prosiect.

Fframwaith Moeseg Data

- 5.32 Mae Fframwaith Moeseg Data (2020) Llywodraeth y DU wedi'i anelu at unrhyw un sy'n gweithio'n uniongyrchol neu'n anuniongyrchol gyda data yn y sector cyhoeddus, gan gynnwys ystadegwyr, gwyddonwyr data, dadansoddwyr a llunwyr polisi. Mae'n 'arwain y defnydd priodol a chyfrifol o ddata yn y llywodraeth ac ... yn helpu gweision cyhoeddus i ddeall ystyriaethau moesegol, mynd i'r afael â'r rhain yn eu prosiectau, ac yn annog arloesi cyfrifol' (Central Digital a Data 2020). Rhennir y fframwaith yn dair egwyddor allweddol gyffredinol: 1) tryloywder, 2) atebolrwydd a 3) tegwch, ac mae'n cynnwys camau gweithredu penodol i asesu a sgorio prosiectau.

Offeryn Hunanasesu Awdurdod Ystadegau'r DU

- 5.33 Nod offeryn hunanasesu moeseg yr UKSA (UKSA 2022) yw 'cynnig fframwaith hawdd ei ddefnyddio i ymchwilwyr adolygu moeseg eu prosiectau drwy gydol y cylch ymchwil'. Mae'n ofynnol i ddadansoddwyr ac ystadegwyr yn yr SYG neu sy'n gweithio mewn llywodraeth fel rhan o Wasanaeth Ystadegol y Llywodraeth (GSS) a Swyddogaeth Dadansoddi'r Llywodraeth ddefnyddio'r offeryn i ystyried y materion moesegol yn eu prosiectau. Mae'r offeryn ei hun yn ddalen Excel sy'n nodi'r agweddau y dylai ymchwilwyr eu hystyried wrth hunanasesu prosiectau yn unol ag egwyddorion moesegol UKSA a'u sgorio yn unol â hynny. Yr egwyddorion yw: lles y cyhoedd, cyfrinachedd a diogelwch data, dulliau ac ansawdd, cydymffurfiaeth gyfreithiol, safbwyntiau ac ymgysylltiad y cyhoedd, a thryloywder. Mae'r offeryn wedi'i gynllunio i ddarparu cysondeb ac i amlygu risgiau moesegol. Mae'r ffocws nid yn unig ar yr hyn y gellir ei wneud gyda data, ond yr hyn y dylid ei wneud a sut y dylid ei wneud.

- 5.34 Os oes gan ddadansoddwyr gynigion prosiect sy'n codi materion moesegol penodol neu'n defnyddio data mewn ffyrdd arloesol neu newydd, gallant ymgynghori â Phwyllgor Cynghori ar Foeseg Data yr Ystadegydd Gwladol (NSDEC) am gyngor. Darperir arweiniad pellach hefyd gan Ganolfan Moeseg Data Cymhwysol yr UKSA ar ystod o bynciau, gan gynnwys lles y cyhoedd a safbwyntiau ac ymgysylltiad y cyhoedd.

The Magenta Book [Y Llyfr Magenta]

- 5.35 Mae'r Magenta Book (Trysorlys EF 2020) yn darparu canllawiau manwl ar werthuso yn y llywodraeth ac mae wedi'i anelu at weithwyr proffesiynol ym meysydd polisi, dadansoddi a chyflawni. Mae'n amlygu bod ystyried moeseg yn broses weithredol a ddylai fod yn rhan o bob cam o'r broses werthuso ac sy'n cynnwys pwyso a mesur buddiannau sy'n cystadlu â'i gilydd. Mae'n cymeradwyo egwyddorion moesegol GSR: budd cyhoeddus, dulliau ymchwil cadarn, cadw at ddeddfwriaeth diogelu data a diogelwch data, caniatâd gwybodus, cyfranogiad a lleihau niwed (GSR 2021).

The Aqua Book [Y Llyfr Aqua]

- 5.36 Mae'r Aqua Book (Trysorlys EM 2015) yn rhoi arweiniad ar sicrhau dadansoddi ansawdd yn y llywodraeth ac yn ymdrin â phynciau fel: asesu addasrwydd i'r diben, ansicrwydd, sicrwydd ansawdd dadansoddol a gweithgareddau gwirio a dilysu. Ni roddir sylw uniongyrchol i foeseg data, er bod sicrhau bod prosesau sicrhau ansawdd effeithiol a bod tryloywder ynghylch ansicrwydd a chyfyngiadau yn ystyriaethau moesegol.
- 5.37 Mae'r fframweithiau a'r dogfennau canllaw uchod yn amlygu'r angen i asesu prosiectau ymchwil yn systematig yn unol â meini prawf ac egwyddorion moesegol allweddol i sicrhau eu bod yn cydymffurfio â'r gyfraith, bod ganddynt fesurau diogelwch data ar waith, eu bod o ansawdd digonol, yn defnyddio dulliau cadarn ac yn nodi risgiau a chyfyngiadau'r ymchwil. Yn ogystal â'r gofynion moesegol sylfaenol hyn, mae hefyd angen asesu prosiectau'n fwy cyfannol i ddeall effeithiau cadarnhaol a negyddol a sicrhau bod egwyddorion moesegol allweddol yn cael sylw ystyrlon. Mae tair o'r egwyddorion hynny, sef tegwch, atebolrwydd, a thryloywder (fel yr amlygir yn y Fframwaith Moeseg Data (Y Swyddfa Ddata a Digidol Ganolog 2020)) yn themâu pwysig yn y llenyddiaeth ac fe'u harchwilir isod.

Llywodraethu da

- 5.38 Mae strwythurau llywodraethu wedi'u hamlygu yn y llenyddiaeth fel rhan bwysig o foeseg data ac mae gwahanol agweddau ar hyn, gan gynnwys tegwch, atebolrwydd a thryloywder. Mae llywodraethu data yn rhan o'r seilwaith llywodraethu cyffredinol ac fe'i diffinnir gan Leonelli (2019) fel 'y strategaethau a'r offer a ddefnyddir i nodi, rheoli a lledaenu data'. Mae'n allweddol i ddehongli data yn ogystal â'r gallu i sicrhau bod modd craffu'n gyhoeddus (Leonelli 2019). Mae pwysigrwydd llywodraethu da mewn mentrau data wedi'i grynhoi gan y Gymdeithas Frenhinol (2020):

Mae llywodraethu da yn dibynnu ar sefydliadau dibynadwy y mae gan y cyhoedd hyder y byddant yn caniatáu i ddata gael eu defnyddio er budd y cyhoedd ond na chaiff eu defnyddio yn eu herbyn. Nid yw'n deg nac yn gredadwy rhoi cyfrifoldeb llywodraethu data moesegol ar unigolion yn unig. Nid yw'n bosibl disgwyl i unigolion bob amser allu penderfynu a yw pob defnydd arfaethedig o ddata amdanynt yn fuddiol neu'n niweidiol (Y Gymdeithas Frenhinol 2020).

Tegwch

- 5.39 Mae tegwch yn thema sy'n treiddio drwy'r cysyniadau moesegol eraill sydd wedi'u cynnwys yn yr adolygiad hwn. Er enghraifft, mae tryloywder 'yn sylfaenol gysylltiedig â thegwch' fel y nodwyd gan yr ICO yn ei diffiniad o dryloywder yn y canllawiau ar-lein ar egwyddorion cyfreithlondeb, tegwch a thryloywder Rheoliad Diogelu Data Cyffredinol y DU (UK GDPR) (ICO 2025).
- 5.40 Mae Kennedy et al. (2020) yn honni bod barn eang y gallai 'arferion data fod yn decach' gyda 'mwy o onestrwydd, tryloywder a deialog gwirioneddol gyda'r cyhoedd.' Wrth ddeall beth mae tegwch yn ei olygu, Mae Deall Data Cleifion a Sefydliad Ada Lovelace (2020) yn amlygu bod tegwch yn gofyn am lywodraethu da, atebolrwydd cyhoeddus a thryloywder. Ymhellach mae'n amlygu bod 'pobl yn malio am ddata'r GIG ac y dylen nhw allu darganfod sut maent yn cael eu defnyddio. Nid yw'n glir i'r cyhoedd sut y gwneir penderfyniadau nawr' (Deall Data Cleifion a Sefydliad Ada Lovelace 2020). Mae'n argymhell y dylai penderfyniadau ynghylch sut y defnyddir data fynd drwy broses dryloyw a bod yn destun goruchwyliaeth allanol (Deall Data Cleifion a Sefydliad Ada Lovelace 2020).

Atebolrwydd

- 5.41 Mae'r Fframwaith Moeseg Data (2020) yn diffinio atebolrwydd yn fras fel y 'mecanweithiau llywodraethu a goruchwyllo effeithiol ar gyfer unrhyw brosiect' (Swyddfa Ddata a Digidol Ganolog 2020). Yn ogystal â hyn, mae'r fframwaith yn nodi bod 'atebolrwydd cyhoeddus' yn golygu bod 'y cyhoedd neu ei gynrychiolwyr yn gallu arfer goruchwyliaeth a rheolaeth effeithiol dros y penderfyniadau a'r camau a gymerir gan y llywodraeth a'i swyddogion...' (Swyddfa Data a Digidol Canolog 2020). Mae Kalkman et al. (2019) yn honni bod 'atebolrwydd cymdeithasol' yn deillio o ymgysylltu ag unigolion mewn cymdeithas.
- 5.42 Mae adroddiad gan Gyngor Biofoeseg Nuffield (2015) ar faterion moesegol casglu a chysylltu data mewn ymchwil biofeddygol a gofal iechyd, yn ymdrin â materion sy'n gysylltiedig â defnydd data mawr a mentrau data yn y sector iechyd, gan gynnwys data gweinyddol a chysylltu data. Mae'n argymhell bod cyrff perthnasol yn deall y niwed posibl y gall camddefnyddio data ei achosi, yn ogystal â'r manteision. Mae'r egwyddorion moesegol ar gyfer mentrau data a ddatblygwyd gan y Cyngor yn nodi 'cyfrifo am benderfyniadau' fel un o'i bedair egwyddor foesegol (Cyngor Biofoeseg Nuffield 2015). Wrth amlinellu beth mae hyn yn ei olygu wrth gyfrifo, mae'r adroddiad yn amlinellu:
- Dylai menter data fod yn destun systemau llywodraethu ac atebolrwydd effeithiol y mae cyfiawnhad moesol drostynt eu hunain. Dylai hyn gynnwys strwythurau atebolrwydd sy'n galw am awdurdod barnwrol a gwleidyddol cyfreithlon, ac atebolrwydd cymdeithasol sy'n deillio o ymgysylltu â phobl mewn cymdeithas. Rhaid i gynnal atebolrwydd effeithiol gynnwys mesurau effeithiol ar gyfer cyfleu disgwyliadau a methiannau llywodraethu, gweithredu a rheolaeth i bobl yr effeithir arnynt ac i'r gymdeithas yn ehangach (Cyngor Biofoeseg Nuffield 2015).
- 5.43 Mae'r llenyddiaeth hon felly'n nodi dwy agwedd wahanol ar atebolrwydd, y gyntaf yn ymwneud â strwythurau a phrosesau i sicrhau bod rhwystrau a gwrthbwysau wedi'u cynnwys mewn prosiectau ymchwil data, a'r ail yn ymwneud â mecanweithiau ar gyfer craffu cyhoeddus drwy ymgysylltu a chyfathrebu effeithiol â'r cyhoedd. Mae adroddiad Cyngor Biofoeseg Nuffield yn galw'r gyntaf yn 'atebolrwydd ffurfiol' a'r ail, yn 'atebolrwydd i'r llw ehangach o randdeiliaid moesol', sef math o atebolrwydd cymdeithasol (Cyngor Biofoeseg

Nuffield 2015). Mae'r adroddiad yn pwysleisio, er mwyn i'r ddwy agwedd ar atebolrwydd fod yn effeithiol, fod angen i bobl gael ffyrdd priodol o ddysgu am yr hyn sydd wedi digwydd i'w data (Cyngor Biofoeseg Nuffield 2015), a bod hyn yn gofyn am 'ystyriaeth ofalus o ddyluniad nid yn unig y systemau y mae menter data benodol yn galw i fod, ond y strwythurau sefydliadol y maent wedi'u hymgorffori ynddynt, er mwyn sicrhau bod digon o dryloywder a chymhellion i roi gwybod am achosion o gam-drin y Cyngor ac i'w hunioni' (Nuffield Council on Bioethics 2015).

- 5.44 Mae'r gallu i geisio iawn yn agwedd bellach ar atebolrwydd ac ymdrinnir ag ef gan adroddiad Cyngor Biofoeseg Nuffield ac yn yr adroddiad Living with Data (Kennedy et al. 2020), a adolygodd dystiolaeth bresennol ar ddealltwriaeth a chanfyddiadau'r cyhoedd o arferion data. Mae'n bosibl bod camddefnydd o ddata yn fwy eang na'r hyn a gydnabyddir gan systemau rheoleiddio. Maent yn debygol o gael eu tangofnodi ac i'r rhai yr effeithir arnynt, mae'n anodd cael iawn (Kennedy et al. 2020).
- 5.45 Kennedy et al. (2020) yn honni bod gwneud iawn yn 'elfen bwysig o ecosystem data teg' oherwydd ei fod yn fodd o fynd i'r afael ag annhegwch. Maent yn dod i'r casgliad o'u hadolygiad o ymchwil bod anfodlonrwydd yn y ffordd y caiff data eu defnyddio a'u rheoli ac yn awgrymu newidiadau sy'n cynnwys 'bodolaeth mesurau diogelu ac atebolrwydd, a'r hawl i wneud iawn; a mwy o reolaeth bersonol' (Kennedy et al. 2020). Mae ymchwil gan Aitkin et al. (2016) wedi dangos bod y cyhoedd yn gwerthfawrogi mwy o reolaeth dros eu data eu hunain o ran pa ddata sy'n cael eu casglu, pwy sydd â mynediad at hyn, sut a gyda phwy y caiff eu rhannu a dibenion eu defnyddio. Cyfeiriwyd at reolaeth yn y modd hwn fel hawl ddynol (Aitkin et al. 2016).
- 5.46 I grynhoi, mae'r llenyddiaeth yn nodi bod angen i atebolrwydd ymgorffori nifer o agweddau nad ydynt wedi'u cyfyngu i brosesau goruchwyllo ffurfiol. Mae'n cynnwys bod yn atebol i'r gymdeithas ehangach ac unigolion drwy ymgysylltu ac esboniad o'r hyn sy'n digwydd i'w data a pham, fel y gall pobl arfer lefel o reolaeth.

Tryloywder

- 5.47 Mae tryloywder yn nodwedd arwyddocaol yn y llenyddiaeth ac yn perthyn yn agos i themâu canolog eraill yn y llenyddiaeth (gan gynnwys y gallu i asesu a yw ymchwil gan ddefnyddio data gweinyddol er budd y cyhoedd, tegwch ac ymddiriedaeth). Mae gwahanol ddiffiniadau o dryloywder yn y llenyddiaeth a chydabyddir llawer o'r heriau o sicrhau tryloywder.
- 5.48 Mae'r Fframwaith Moeseg Data (2020) yn nodi: 'Mae tryloywder yn golygu bod eich gweithredoedd, eich prosesau a'ch data yn agored i'w harchwilio trwy gyhoeddi gwybodaeth am y prosiect mewn fformat cyflawn, agored, dealladwy, hygyrch, a rhad ac am ddim.' (Swyddfa Ddigidol a Data Ganolog 2020). Mae llenyddiaeth Sefydliad Ada Lovelace (2021 a 2022) yn cyfeirio at 'dryloywder ystyrion' ac yn gwahaniaethu rhwng 'tryloywder' ac 'eglwurhad'. Eglurhad yng nghyd-destun Deallusrwydd Artiffisial yw 'i ba raddau y gellir esbonio mecaneg fewnol peiriant neu system ddysgu dwfn neu ei gwneud yn ddealladwy mewn termau dynol' (Sefydliad Ada Lovelace 2021).
- 5.49 Bydd manylder tryloywder yn dibynnu ar y math o ddata a ddefnyddir, pwrpas eu defnyddio a'r gynulleidfa. Mae'r awduron yn honni, er mwyn i dryloywder fod yn ystyrion, bod yn rhaid iddo symud y tu hwnt i drosglwyddiad syml o wybodaeth i wrthrychau data ynghylch sut y defnyddir eu data. Dylai ganiatáu 'buddiolwyr i gael cipolwg ar y prosesau gwneud penderfyniadau sy'n pennu eu nodau, effaith eu canlyniadau, sut maent yn rhyngweithio â strwythurau cymdeithasol a pha gysylltiadau pŵer y maent yn eu meithrin trwy ddefnydd parhaus' (Sefydliad Ada Lovelace 2021). At hynny, mae'n cyfeirio at fentrau data a phrosiectau sy'n 'adeiladu cyhoeddusrwydd' ac sy'n rhoi digon o ddealltwriaeth a gwybodaeth i bobl rhyngweithio â strwythurau data neu gymryd camau os dymunant (Sefydliad Ada Lovelace 2021).
- 5.50 Canfu ymchwil ansoddol gan y cwmni ymchwil Britain Thinks (Thinks Insight and Strategy erbyn hyn) i'r hyn y mae'r cyhoedd yn ei feddwl am dryloywder (wrth wneud penderfyniadau algorithmig) yn y sector cyhoeddus nad oes gan bobl 'gysylltiadau pen meddwl cryf' â'r cysyniad o dryloywder a bod dryswch ynghylch yr hyn y mae'n ei olygu, gan nad yw'n rhywbeth y mae llawer o bobl yn ymgysylltu ag ef fel mater o drefn (Britain Thinks 2021).

- 5.51 Mae un enghraifft o'r llenyddiaeth, er nad yw'n ymwneud yn uniongyrchol ag ymchwil data gweinyddol, yn dangos sut y gellir sicrhau tryloywder. Mae'r 'gofrestr algorithm' a weithredwyd gan ddinasoedd Amsterdam a Helsinki yn dangos sut y defnyddir systemau data a deallusrwydd artifisial yng ngwasanaethau cyhoeddus y ddinas ac maent ar gael i unrhyw un ymgynghori â hwy. Maent yn darparu gwahanol lefelau o wybodaeth ar gyfer gwahanol gynulleidfaoedd, ac maent yn amlinellu ffeithiau allweddol am systemau a yrrir gan ddata ac yn caniatáu i bobl wirio sut y gall y systemau hyn effeithio arnynt (Ada Lovelace Institute 2021).
- 5.52 Mae'r llenyddiaeth yn amlinellu manteision tryloywder. Mae'n helpu i hwyluso ymddiriedaeth yn y llywodraeth a meithrin cefnogaeth y cyhoedd i brosiectau data. Yng nghyd-destun ymchwil data gweinyddol, ceisiodd adolygiad llenyddiaeth ADR UK ar ymddiriedaeth, diogelwch a budd y cyhoedd 'archwilio agweddau'r cyhoedd, fel y canfuwyd gan ymchwil flaenorol, tuag at rannu, cysylltu a defnyddio data gweinyddol ar gyfer ymchwil, a'r amodau y canfyddir y dylai ac na ddylai ddigwydd odanynt.' (ADR UK 2020). Canfu fod tryloywder yn effeithio'n uniongyrchol ar gefnogaeth y cyhoedd ar gyfer ymchwil gan ddefnyddio data gweinyddol a bod 'ymddiriedaeth a thryloywder gyda'i gilydd felly yn ffurfio amod allweddol arall o gefnogaeth y cyhoedd i ymchwil sy'n defnyddio data gweinyddol, gyda thryloywder yn cael effaith sylweddol ar lefelau ymddiriedaeth'. (ADR UK 2020). Yn ogystal, mae bodolaeth tryloywder yn bwysig ar gyfer gwneud y mwyaf o fudd cyhoeddus ymchwil gymdeithasol fel y cydnabyddir gan ganllawiau sicrwydd moesegol GSR:
- Mae sicrhau bod dulliau a chanfyddiadau ymchwil yn cael eu lledaenu'n dryloyw yn gwneud y mwyaf o fudd cyhoeddus posibl ymchwil gymdeithasol, trwy alluogi canfyddiadau'r ymchwil i gyrraedd cynulleidfaoedd ehangach. Mae hefyd yn rhan hanfodol o gynnal ymddiriedaeth y cyhoedd wrth gasglu a defnyddio data cyhoeddus (GSR 2021).
- 5.53 Mae'r llenyddiaeth hefyd yn amlygu cymhlethdod tryloywder. Mae ymchwil Kennedy et al. (2022) yn nodi y gall defnyddiau data fod yn anodd eu deall, a rhagdybir y bydd mwy o dryloywder yn gwella dealltwriaeth ac y bydd hyn yn arwain at fwy o ymddiriedaeth ac agweddau cadarnhaol ynghylch defnyddiau data. Mae Kennedy et al. 'yn dangos bod angen gwell dealltwriaeth arnom o'r hyn y mae'n ei olygu i ddeall defnyddiau data' oherwydd nid yw gwybodaeth glir

o reidrwydd yn arwain at well dealltwriaeth. At hynny, nid yw cyfathrebu a bod yn dryloyw yn unig yn ddigon, ond bod angen i gyfathrebu fynd i'r afael â phryderon pobl. Yn ogystal, yr hyn sydd ei angen fwyaf yn ôl Kennedy et al. yw 'defnyddiau data gwell' fel eu bod yn 'dileu niwed a'u bod er budd y cyhoedd neu gymdeithas' (Kennedy et al. 2022).

- 5.54 Mae angen sicrhau cydbwysedd rhwng rhoi gormod o wybodaeth, a all gynrychioli risg a rhoi rhy ychydig, a all erydu ymddiriedaeth a chyfyngu ar gyfranogiad. Fe wnaeth Kennedy et al. (2022) ganfod mai'r bobl sy'n gwybod fwyaf am ddefnyddiau data sy'n aml yn poeni fwyaf amdanynt. Mae Sexton et al yn cydnabod rhai o'r heriau:

I mi, fel eiriolwr tryloywder, mae hyn yn hynod o heriol i mi oherwydd rwyf am fod allan yn siarad â'r cyhoedd am yr hyn rydym yn ei wneud gyda'u data, a'u cael i roi eu barn ...ond mae rhannu data a chysylltu data yn syniadau hynod gymhleth mewn gwirionedd. ...rwy'n meddwl nad ydym wedi bod yn ddigon da am ddangos i bobl sut olwg sydd ar eu gwybodaeth, a fy mhrofiad i yw eich bod yn cael eich gwrthio i wneud 20 munud gyda grŵp o 15 o gleifion yn ceisio dweud wrthynt beth mae'r rhaglen yn ei wneud a'u cael i ddweud 'ie, mae'n swnio fel syniad da'. A dweud y gwir, mae'n anodd iawn gwneud hynny ... ond os ydych chi am iddyn nhw gyfrannu'n iawn at yr hyn rydym yn ei wneud mae'n rhaid i chi dreulio amser yn eu haddysgu ac yn adeiladu'r sylfaen wybodaeth honno cyn iddyn nhw allu dechrau mewnbynnu'n iawn. (Sexton et al. 2017).

- 5.55 Roedd Sexton et al. (2017) yn mynegi'r tensiynau sy'n bodoli oherwydd y diffyg gwybodaeth mewn defnyddiau data cymhleth a'r mater ehangach o adeiladu gwybodaeth gyhoeddus yn y meysydd hyn cyn y gellir cyflawni tryloywder a chyfranogiad. Rhoi gwybodaeth hygyrch glir yw'r cam cyntaf i gyfranogiad ehangach dinasyddion (Sefydliad Ada Lovelace 2021) ac mae hyn yn golygu rhannu gwybodaeth 'am yr hyn sy'n digwydd i ddata pobl mewn ffyrdd sy'n eu galluogi i weithredu' (Sefydliad Ada Lovelace 2021). Mae hyn yn awgrymu bod y meincnod ar gyfer tryloywder yn ymwneud â ph'un a all pobl ddefnyddio'r gallu i ddefnyddio eu data ac mae rhan o hyn yn ymwneud â galluogi dealltwriaeth a sgil.

- 5.56 O ran sut y gellir cyflawni hyn, mae adroddiad Britain Thinks (2021) yn dangos bod awydd cyffredin i wybodaeth dryloywder (boed yn sylfaenol neu'n fwy

cymhleth) gael ei chyflwyno'n glir ac yn syml. Mae tryloywder yn benodol i'r cydestun gyda chyfathrebu mwy gweithredol yn cael ei ystyried yn fwy priodol ar gyfer achosion risg uwch effaith uwch. Gallai rhai o'r problemau gael eu datrys drwy gael haenau gwahanol o wybodaeth gyda mwy o wybodaeth ar gael i bobl yr effeithir arnynt yn uniongyrchol, neu sydd am gael gwybod mwy (Britain Thinks 2021).

- 5.57 At hynny, fel yr amlygwyd yn adroddiad Data Analytics Research Environments UK (DARE UK) ar adeiladu seilwaith ymchwil data cenedlaethol dibynadwy, mae angen i geidwaid data ac ymchwilwyr fod yn fwy rhagweithiol o ran bod yn dryloyw. Gallent ddefnyddio ystod o sianeli hygyrch i gyflawni hyn. Er enghraifft, 'Dylid dod â gwybodaeth yn rhagweithiol i fywydau pobl trwy ddulliau cymysg: er enghraifft, trwy arweinwyr cymunedol y gellir ymddiried ynddynt fel meddygon teulu, cynghorau lleol, a sefydliadau ffydd, a thrwy ymgyrchoedd cyfryngau cymdeithasol a hysbysebu, addysgu mewn ysgolion, ar hysbysfyrddau cymunedol, a thrwy siarad â phobl a dosbarthu gwybodaeth mewn mannau cyhoeddus' (Harkness et al. 2022).
- 5.58 I grynhoi, mae'r llenyddiaeth yn amlinellu bod llawer o fanteision i ddarparu gwybodaeth i'r cyhoedd yn ehangach am brosiectau ymchwil data sy'n dibynnu ar gefnogaeth y cyhoedd. Fodd bynnag, nid yw darparu gwybodaeth yn unig yn ddigon, mae sut y gwneir hyn yn bwysig ac mae angen gwell dealltwriaeth ac ymwybyddiaeth gyhoeddus o gysyniadau a defnyddiau data. Mae angen i bobl wybod beth sy'n digwydd i'w data er mwyn caniatáu iddynt gymryd rhan mewn deialog am ymchwil data gweinyddol (Sefydliad Ada Lovelace 2021).

Budd y cyhoedd

- 5.59 Mae cynnal ymchwil data er budd y cyhoedd yn thema ganolog yn y llenyddiaeth foeseg. Fe'i nodir fel egwyddor foesegol yng Nghanllawiau Proffesiynol GSR (GSR 2021), y Fframwaith Moeseg Data (Swyddfa Ddigidol a Data Ganolog 2020) a chan yr UKSA. Defnyddir termau gwahanol megis 'lles y cyhoedd', budd y cyhoedd' a 'budd cyhoeddus' ac mae hyn yn ychwanegu lefel o ansicrwydd ynghylch ei ystyr a'i gymhwysiad. At ddibenion yr adolygiad hwn defnyddir 'lles y cyhoedd' yn bennaf, ond cynhwysir llenyddiaeth sy'n defnyddio termau eraill gan eu bod yn aml yn cael eu defnyddio'n gyfnewidiol.

- 5.60 Mae llawer o'r llenyddiaeth a'r ymchwil yn y maes hwn yn ceisio diffinio ystyr budd y cyhoedd. Oherwydd y lluosogrwydd o dermau a hyblygrwydd y dehongliad y gellir eu rhoi iddynt, mae Laurie a Stevens (2016) yn amlygu bod 'budd y cyhoedd' yn derm sy'n llawn hyblygrwydd dehongliadol'. Maent yn amlygu nad oes modd dweud sut y mae'n berthnasol ar draws yr amrywiaeth o ffrydiau gwaith ar draws y sector cyhoeddus a bod yn rhaid i'w hystyr gael ei ddiffinio gan awdurdodau cyhoeddus eu hunain yn ôl y cyd-destun. (Laurie a Stevens 2016). Mae llenyddiaeth arall yn nodi bod yn rhaid penderfynu arno fesul achos (Harkness et al. 2022).
- 5.61 O dan y pŵer ymchwil yn Neddf yr Economi Ddigidol (DEA) 2017, gellir rhannu gwybodaeth a gedwir gan awdurdodau cyhoeddus at ddibenion ymchwil. Nod y Ddeddf yw rhoi sicrwydd ac eglurder i gyrff cyhoeddus sy'n datgelu data ar gyfer ymchwil. Mae'r Cod Ymarfer Ymchwil a Meini Prawf Achredu (DEA 2017) yn nodi'r amodau ar gyfer gwneud hyn. Un amod o'r fath yw bod yr ymchwil er budd y cyhoedd. Mae'r Cod Ymarfer yn rhoi enghreifftiau o brosiectau er budd y cyhoedd fel ymchwil sydd â'r prif ddiben o:
- darparu sylfaen dystiolaeth ar gyfer gwneud penderfyniadau polisi cyhoeddus
 - darparu sylfaen dystiolaeth ar gyfer darparu gwasanaethau cyhoeddus
 - darparu sylfaen dystiolaeth ar gyfer penderfyniadau sy'n debygol o fod o fudd sylweddol i'r economi, cymdeithas, neu ansawdd bywyd pobl yn y DU
 - atgynhyrchu, dilysu, herio neu adolygu ymchwil presennol a chyhoeddiadau ymchwil arfaethedig, gan gynnwys ystadegau swyddogol
 - ymestyn dealltwriaeth o dueddiadau neu ddigwyddiadau cymdeithasol neu economaidd yn sylweddol trwy wella gwybodaeth neu herio dadansoddiadau a dderbynnir yn eang
 - gwella ansawdd, cwmpas neu gyflwyniad ymchwil presennol, gan gynnwys Ystadegau swyddogol neu Genedlaethol (DEA 2017).
- 5.62 Mae'r Cod Ymarfer yn fan cychwyn defnyddiol ar gyfer deall budd y cyhoedd, ac ategir hyn gan ganllawiau a llenyddiaeth arall o ran sut y gellir ei asesu mewn prosiectau ymchwil. Mae canllawiau moeseg yr UKSA (ar ystyried lles y cyhoedd mewn ymchwil ac ystadegau) yn esbonio lles y cyhoedd trwy gyfeirio

at y ddeddfwriaeth hon ac yn crynhoi fel a ganlyn: 'Yn y bôn, mae lles y cyhoedd yn ymwneud â buddion eich ymchwil a'ch dadansoddiad i'r cyhoedd: Pa effaith a gaiff eich gwaith ar bobl, cymunedau, sefydliadau, a chwmnïau?' (Canolfan Moeseg Data Cymhwysol UKSA 2021).

- 5.63 Mae'r canllawiau yn amlygu pwysigrwydd deall y math o effaith, maint yr effaith, cwmpas y boblogaeth a strategaeth effaith yr ymchwil (h.y. y camau a gymerwyd i sicrhau bod y cyhoedd yn gweld y canfyddiadau). Wrth asesu hyn mae hefyd yn bwysig meddwl am risgiau ac mae'r UKSA yn darparu rhestr wirio i gydbwysu risgiau a buddion prosiectau ymchwil i nodi niwed, ffactorau lliniarol a chyfiawnhad. Yn gyffredinol, fodd bynnag, mae'n cydnabod pwysigrwydd cyddestun a rhesymeg y gwaith a bod 'asesu lles y cyhoedd yn broses gynhenid oddrychol' (UKSA Centre for Applied Data Ethics 2021). Fodd bynnag, efallai nad yw hyn yn syml oherwydd, fel yr amlygwyd gan Scott et al. 2018, nid oes unrhyw ganllawiau na fframwaith i awdurdodau asesu neu werthuso budd cyhoeddus.
- 5.64 Darperir diffiniadau eraill hefyd yn y llenyddiaeth. Er enghraifft, mae Jones a Ford (2018) yn diffinio 'budd cyhoeddus' fel gwaith sydd â gwerth yn y byd go iawn neu gymwysiadau ymarferol gyda'r potensial clir i wella llawer o unigolion a/neu gymdeithas ehangach' (Jones a Ford 2018).
- 5.65 Mae'r Fframwaith Moeseg Data (Swyddfa Ddigidol a Data Ganolog 2020) yn cynghori bod y cwestiynau canlynol yn cael eu gofyn i fentrau a phrosiectau data i helpu i bennu'r budd cyhoeddus ehangach yn eu cyd-destun:
- Beth yw'r manteision uniongyrchol i unigolion yn y prosiect hwn?
 - Sut mae'r prosiect yn sicrhau canlyniadau cymdeithasol cadarnhaol i'r cyhoedd ehangach?
 - Sut gallwch chi fesur a chyfleu manteision y prosiect hwn i'r cyhoedd?
 - Beth yw'r grwpiau a fyddai dan anfantais neu na fyddai'n elwa o'r prosiect? (Swyddfa Ddigidol a Data Ganolog 2020)
- 5.66 Mae nifer o ymarferion deialog cyhoeddus wedi archwilio sut mae'r cyhoedd yn deall 'budd cyhoeddus' mewn ymchwil. Fel y nodwyd uchod, ceisiodd ADR UK a'r OSR ddeall beth mae 'lles y cyhoedd' yn ei olygu mewn ymchwil ac ystadegau trwy ddeialog cyhoeddus ledled y DU (ADR UK ac OSR 2022). Rhai

o'r canfyddiadau allweddol oedd bod y cyhoedd am gael eu cynnwys mewn penderfyniadau ynghylch a yw lles y cyhoedd yn cael ei wasanaethu, a dylai ymchwil fynd i'r afael ag anghenion y byd go iawn a mynd i'r afael ag anghydraddoldeb cymdeithasol. Dylid hefyd fod cyfathrebu clir, rhagweithiol i godi ymwybyddiaeth o'r defnydd o ddata (Rheoliad y Swyddfa Ystadegau 2022).

5.67 Nod Adroddiad y Gwarcheidwad Data Cenedlaethol (2022) oedd diffinio budd cyhoeddus trwy broses o ddeialog cyhoeddus. Canfu fod angen rhywfaint o 'les net' o'r gwaith sydd ag agwedd 'budd' ac agwedd 'gyhoeddus'. Mae'r agwedd budd yn gofyn am gyflawni lles uniongyrchol, anuniongyrchol, brys neu hirdymor, wedi'i gydbwyso yn erbyn risg a rhaid iddo fod yn adnabyddadwy. Mewn perthynas â gweithgareddau masnachol, canfu'r adroddiad os mai elw yw'r unig fudd, yna nid yw'r defnydd er budd y cyhoedd (Gwarcheidwad Data Cenedlaethol 2022).

5.68 Mae rhai sylwebwyr yn feirniadol o'r modd y mae gan 'fudd y cyhoedd' y potensial i gael ei drin yn unol â'r prif fuddiannau. Roedd Muller et al. (2021) yn wylidwrs o'r naratif ynghylch datblygu trwydded gymdeithasol ar gyfer defnyddio data mewn ymchwil er budd y cyhoedd gan yr honnir mai'r fframwaith hwn yw:

...wedi'i amgylchynu'n fawr gan bersbectif y nodau ymchwil trech, a bennwyd ymlaen llaw, sy'n ddwys o ran data.... Mae llywodraethu yn cael ei drwytho â normau sy'n ysgogi dinasyddion, neu'n eu 'hannog', i ymddwyn yn unol â barn offerynnol a bennwyd ymlaen llaw o fudd y cyhoedd. O fewn y farn hon, mae mewnbwn o'r gwaelod i fyny yn cael ei weld fel offeryn defnyddiol ar gyfer dylanwadu ar ymddygiad. Ac felly nid yw budd y cyhoedd yn cael ei weld fel nod ynddo'i hun, ond yn hytrach fel ffordd tuag at y diwedd o wneud ymchwil iechyd sy'n defnyddio mwy o ddata yn bosibl (Muller et al. 2021).

5.69 Mae Sefydliad Ada Lovelace (2022) hefyd yn hollbwysig, ac mae'n gofyn: Pwy sy'n rhan o'r sgwrs ynghylch pennu defnydd data er budd y cyhoedd? Pwy sy'n penderfynu beth yw 'lles cyhoeddus' neu 'fudd cymdeithasol', a sut y gellir gwneud penderfyniadau o'r fath yn gyfiawn? (Sefydliad Ada Lovelace 2022). Mae'n cwestiynu a yw'r penderfyniadau hyn yn cael eu gwneud mewn ffordd gynhwysol a chyfranogol ac yn ystyried: 'Mae'r syniad o 'les y cyhoedd' yn agor cwestiynau pwysig ynghylch pa fath o 'cyhoedd' sy'n rhan o drafodaethau a

phwy sy'n cael eu gadael allan.' (Sefydliad Ada Lovelace 2022). Yn wir, fel yr amlygwyd gan Kennedy et al. 2022, ceir amrywiaeth o gyhoeddusrwydd gan fod gan wahanol bobl farn wahanol am ddefnyddiau data ac mae'n bwysig 'ymgynggori'n rheolaidd â chyhoedd amrywiol, oherwydd bod defnyddiau data a chanfyddiadau ohonynt yn newid, ac oherwydd bod anghydraddoldebau strwythurol yn dylanwadu ar yr hyn y mae gwahanol bobl yn ei feddwl am wahanol ddefnyddiau data' (Kennedy et al. 2022).

- 5.70 I grynhoi, gellir gweld o'r llenyddiaeth y gellir diffinio budd y cyhoedd mewn gwahanol ffyrdd oherwydd ei fod yn benodol i'r cyd-destun. Fodd bynnag, mae'r pwyntiau allweddol i'w hystyried wrth asesu budd y cyhoedd yn cynnwys a yw'r ymchwil yn ychwanegu at y sylfaen dystiolaeth i helpu i wneud penderfyniadau, neu a yw'n cyfrannu at ymchwil neu wybodaeth sy'n bodoli eisoes. Gall fod yn bwysig nodi agwedd 'budd' ac 'cyhoeddus' i'r ymchwil a mynd trwy ymarfer sy'n cydbwysu effeithiau, buddion, risgiau neu niwed posibl, yn ogystal ag anghydraddoldeb strwythurol. Mae angen cyfleu budd y cyhoedd yn eang a bod yn gynhwysol mewn trafodaethau am yr hyn y mae budd y cyhoedd yn ei olygu. Oherwydd hyblygrwydd y dehongli a natur cyd-destun penodol budd y cyhoedd, gellir ei ddefnyddio gan leisiau pwerus yn y seilwaith data i gyfiawnhau camau gweithredu penodol.

Ymddiriedaeth

- 5.71 Mae'r cysyniad o ymddiriedaeth yn ganolog i lawer o'r themâu a amlygir yn y llenyddiaeth ynghylch defnyddio data gweinyddol yn foesegol mewn ymchwil, gan fod ymchwil o'r fath yn dibynnu ar gefnogaeth y cyhoedd. Canfu ymchwil ADR UK (2020) fod pobl yn deall 'dibynadwyedd' fel ffordd o ddeall os yw ymchwil yn cael ei wneud yn foesegol ac yn deg:

Yn hytrach na defnyddio iaith foesegol ac anfoesegol, roedd y cyfranogwyr yn aml yn dychwelyd i iaith dibynadwyedd. Roedd trafodaethau ynghylch dibynadwyedd yn aml wedi'u gwreiddio yn yr hyn yr oedd cyfranogwyr yn ei ystyried yn arfer da neu foesegol; felly, roedd eu canfyddiad o faterion moesegol yn gyffredinol yn adlewyrchu eu hargymhellion ar gyfer dibynadwyedd (ADR UK 2020).

- 5.72 Mae'r adroddiad yn honni bod ymddiriedaeth a thryloywder yn 'amodau allweddol ar gyfer cefnogaeth y cyhoedd i rannu a chysylltu data gweinyddol ar

gyfer ymchwil' (ADR UK 2020). At hynny, mae'n honni bod ymddiriedaeth yn amrywio yn ôl sefydliad ac 'nad yw pob corff cyhoeddus yn derbyn yr un lefelau o ymddiriedaeth, a bod mwy o ymddiriedaeth mewn rhai sefydliadau nag eraill' (ADR UK 2020).

- 5.73 Mae'r llenyddiaeth hefyd yn dangos bod ymddiriedaeth y cyhoedd yn dod gydag amodau, ei fod yn ddibynnol ar gyd-destun, yn gymhleth ac yn amrywiol fel y nodwyd gan Kennedy et al. (2020):

...mae pobl yn ymddiried mwy mewn rhai sectorau gyda'u data nag eraill. Mae'r berthynas rhwng ymddiriedaeth mewn sefydliadau yn gyffredinol ac ymddiriedaeth yn arferion data sefydliadau yn gymhleth, ac mae canfyddiadau'n gwrth-ddweud ei gilydd (Kennedy et al. 2020).

- 5.74 Mae adroddiad Mynd i'r afael ag Ymddiriedaeth mewn Defnydd Cyhoeddus o Ddata (Canolfan Moeseg Data ac Arloesi 2021) yn canfod bod amgylchedd o ymddiriedaeth denau lle rhennir data at ddibenion pwysig, a bod problemau gyda'r ffordd y caiff hyn ei gyfleu. Mae'n honni ei bod yn debygol y bydd amrywiaeth yn ymddiriedaeth y cyhoedd sy'n benodol i'r cyd-destun. Canfu'r adroddiad hefyd fod cyfranogwyr yn ymwybodol o anghydbwysedd pŵer rhwng perchnogion data a gwrthrych y data, ac (fel testunau data), roeddent yn teimlo ymdeimlad o ddiffyg grym yn y modd y defnyddiwyd eu data gan fod hyn yn rhywbeth i'w benderfynu gan sefydliadau a oedd yn cadw eu data (Canolfan Moeseg Data ac Arloesi 2021).

- 5.75 Mae'n bosibl bod rhan o'r diffyg grym hwn yn ymwneud â natur gymhleth, dechnegol a haniaethol ymchwil data mawr lle mae'r data 'wedi'u hymestyn i ffwrdd o'i gyd-destun gwreiddiol' (Sexton et al. 2017). Fel y nodwyd uchod, mae'n dod yn anodd iawn i unigolion olrhain seilweithiau data mawr, heb sôn am gael goruchwyliaeth ar ansawdd, arwyddocâd gwyddonol ac effaith gymdeithasol bosibl y wybodaeth sy'n cael ei chynhyrchu (Leonelli 2020).

- 5.76 Mae meithrin ymddiriedaeth y cyhoedd yn anodd felly oherwydd fel y nodwyd gan Sexton et al. (2017), 'mae rhannu ac ailddefnyddio data yn yr amgylcheddau hyn yn fwyfwy anodd ei ddeall, ac yn gynyddol heriol yn foesebol. Nid oes ateb hawdd ar sut i chwalu'r cymhlethdod hwn a galluogi trafodaeth gyhoeddus wybodus' (Sexton et al. 2017). Fodd bynnag, mae meithrin deialog cyhoeddus ehangach a chael y cyfathrebu'n gywir drwy fynd i'r

afael â'r materion y mae pobl am eu deall (Kennedy et al. 2022) yn rhan o'r ateb. Yn ogystal, mae'r llenyddiaeth yn amlygu'r angen am gyfathrebu rhagweithiol fel bod gwybodaeth glir yn cael ei dwyn i fywydau pobl trwy ystod o sianelau a dulliau i godi ymwybyddiaeth gyffredinol o arferion ymchwil data fel un ffordd bosibl o feithrin ymddiriedaeth y cyhoedd (Harkness et al. 2022).

Ymgysylltu â'r cyhoedd

- 5.77 Gellir gweld trwy gydol yr adolygiad hwn bod ymgysylltu â'r cyhoedd yn bloc adeiladu sylfaenol mewn ymchwil data moesegol. Mae ymgysylltu â'r cyhoedd yn un o egwyddorion moesegol UKSA ac mae'n rhan o'r Fframwaith Moeseg Data (Swyddfa Ddigidol a Data Ganolog 2020). Mae gwahanol ddiffiniadau o ymgysylltu â'r cyhoedd. Mae ADR UK yn ei ddiffinio ar eu gwefan fel 'set bwrpasol o weithgareddau wedi'u cynllunio i hyrwyddo deialog barhaus gyda'r cyhoedd am ymchwil data gweinyddol, wedi'i sbarduno gan wrando gweithredol ac ymateb.' (gwefan ADR UK).
- 5.78 Cydnabyddir yn gyffredinol yn y llenyddiaeth bod angen mwy o ymgysylltu â'r cyhoedd. Amlygir hyn yn adroddiad yr OSR: Rhannu Data a Chysylltu er Lles y Cyhoedd (2023) sy'n nodi mai ymgysylltu â'r cyhoedd mewn perthynas â rhannu data a chysylltiadau oedd un o'r pynciau mwyaf a grybwyllwyd gan gyfweleion. Amlygodd hefyd y gall fod diffyg dealltwriaeth ynghylch sut i ymgysylltu â'r cyhoedd yn dda. Mae'n mynd ymlaen i 'gytuno â'r angen am ragor o ymgysylltu â'r cyhoedd' a chydabu fod ymgysylltu â'r cyhoedd yn bwysig ar gyfer cynyddu dealltwriaeth ac i 'gynyddu o bosibl faint o drwydded gymdeithasol sydd gan brosiect rhannu data neu gysylltu data (OSR 2023).
- 5.79 Mae'r llenyddiaeth yn ymhelaethu ar ymgysylltu â'r cyhoedd mewn nifer o ffyrdd drwy ei gysylltu â sefydlu 'trwydded gymdeithasol' ar gyfer defnyddio data. I bob pwrpas, mae trwydded gymdeithasol yn ddealltwriaeth neu'n set o ddisgwyliadau rhwng y wladwriaeth a dinasyddion ynghylch sut y bydd data'n cael eu defnyddio. Cyflwynwyd y cysyniad gan y cymdeithasegydd meddygol Pam Carter a chydweithwyr yn 2015 ac fe'i dyfynnwyd gan Muller et al. (2021). Fe'i disgrifir fel 'disgwyliadau cymdeithas o ran ymddygiad a gweithgareddau corfforaethau sy'n mynd y tu hwnt i ofynion rheoleiddio ffurfiol...mae amodau angenrheidiol y cysyniad yn cynnwys dwyochredd, peidio â chamfanteisio a gwasanaeth er lles y cyhoedd' (Muller et al. 2021). Yng nghyd-destun rhannu

data er budd y cyhoedd ehangach, mae'r cysyniad o drwydded gymdeithasol yn bwysig fel y cydnabyddir gan y Ganolfan Moeseg Data ac Arloesedd (2020):

Er bod gennym ddealltwriaeth ddatblygedig o'r contract cymdeithasol rhwng y dinesydd a'r wladwriaeth mewn perthynas â threthiant a gwariant cyhoeddus, nid yw'r un peth yn wir am ddata. O ystyried bod data efallai cyn bwysiced ag arian i weithrediad y wladwriaeth, ac mae ei arwyddocâd yn cynyddu, rydym yn dadlau ei bod yn bryd ystyried contract cymdeithasol clir rhwng y dinesydd a'r wladwriaeth ynghylch sut y caiff data eu rhannu a'u defnyddio (Canolfan Moeseg Data ac Arloesi 2020).

- 5.80 Mae llenyddiaeth Sefydliad Ada Lovelace yn rhagweld defnyddio ystod ehangach o ddulliau ymgysylltu a chyfranogol at ddata fel bod pobl yn cael eu cynrychioli'n well, yn gallu ymgysylltu'n ystyrion a chymryd rhan mewn penderfyniadau am ddata a sut y caiff eu llywodraethu:

Trwy ystod eang o ddulliau cyfranogol - o gynghorau a rheithgorau dinasyddion sy'n llywio polisi a rheoleiddio data lleol a chenedlaethol yn uniongyrchol, i gynrychiolaeth gyhoeddus ar fyrddau llywodraethu cwmnïau technoleg - mae pobl yn cael eu cynrychioli'n well, yn cael eu cefnogi a'u grymuso'n well i wneud i systemau a seilweithiau data weithio iddynt, ac mae llunwyr polisi yn fwy gwybodus am yr hyn y mae pobl yn ei ddisgwyl ac yn ei ddymuno o ddata, technolegau a'u defnyddiau. (Sefydliad Ada Lovelace 2022).

- 5.81 Ymhellach, fel y nodwyd uchod, mae llenyddiaeth Sefydliad Ada Lovelace yn honni bod angen 'stiwardiaeth' ar ddata. Mae hwn yn 'gysyniad sy'n ymgorffori cynllunio cyfrifol a rheoli adnoddau cyffredin' sy'n ymwneud ag ail-gysyniadu data a'u deall fel 'lles cyffredin y mae gan bawb ran ynddo, a lle dylai buddiannau buddiolwyr fod wrth galon pob sgwrs' (Ada Lovelace Institute 2021).

- 5.82 O ran sut y gellir cyflawni ymgysylltiad gwell mewn ymchwil cysylltu data, mae Aitkin et al wedi awgrymu. (2016) y gellid defnyddio dulliau ymchwil ansoddol fel mathau o ymgysylltu â'r cyhoedd a allai ennyn cefnogaeth ac ymddiriedaeth y cyhoedd:

Yn y modd hwn gallai mwy o ddefnydd o ddulliau ansoddol fod yn sylfaen ar gyfer cymorth. Mae ymgysylltu â'r cyhoedd ac ymchwil ansoddol o'r fath yn

elfennau cynyddol aml o brosiectau gwyddoniaeth mawr ac yn cynrychioli, yn rhannol, ymdrechion i gynyddu ymddiriedaeth y cyhoedd a sicrhau ymchwil ac arloesi cyfrifol (Aitkin et al. 2016).

- 5.83 Yn ogystal â defnyddio amrywiaeth o ddulliau, mae adroddiad DARE UK ar adeiladu seilwaith ymchwil data cenedlaethol dibynadwy, yn nodi y dylai ymchwilwyr feddu ar y sgiliau i ymgysylltu â'r cyhoedd mewn ffordd ystyrlon (Harkness et al. 2022). Mae hyn yn amlygu'r cysylltiad rhwng ymgysylltu effeithiol â'r cyhoedd a chanfyddiadau o ymchwil data fel rhywbeth y gellir ymddiried ynddo.
- 5.84 I grynhoi, mae'r llenyddiaeth yn cydnabod bod ymgysylltu â'r cyhoedd yn rhan hanfodol o ymchwil foesegol ac yn gysylltiedig â datblygu trwydded gymdeithasol ar gyfer defnyddio data. Mae angen rhagor o waith i ddad-ddewis y sgiliau, y strwythurau a'r dulliau gweithredu sydd eu hangen i ymgysylltu â'r cyhoedd yn dda. Mae rhywfaint o waith wedi'i wneud i fynd i'r afael â'r materion hyn. Mae Menter Ymgysylltu â'r Cyhoedd mewn Ymchwil Data (PEDRI) yn bartneriaeth sydd newydd ei chreu o sefydliadau sy'n gweithio gyda data, sydd â'r nod o ddatblygu arfer gorau a 'rhannu mewnwelediadau, offer a chanllawiau i bobl sy'n gwneud ymchwil data i weithio'n dda gyda'r cyhoedd' (gwefan PEDRI 2025). Maent wedi datblygu canllawiau ar gyfer ymgysylltu â'r cyhoedd mewn ymchwil data.

6. Casgliadau

- 6.1 Mae'r adolygiad hwn wedi ystyried detholiad o'r llenyddiaeth ddiweddar ar foeseg data sy'n berthnasol i ymchwil data gweinyddol i ddeall ac archwilio'r themâu moesegol mwyaf cyffredin i ymchwilwyr a dadansoddwyr y llywodraeth. Roedd rhan o hyn yn ymwneud â llenyddiaeth ar gyd-destun ehangach digideiddio cynyddol mewn cymdeithas. Felly cyflawnodd ei nodau i archwilio'r sylfaen dystiolaeth ddiweddar a nodi fframweithiau a themâu moesegol perthnasol. Mae argymhellion ar arfer gorau sy'n seiliedig ar y dystiolaeth yn y llenyddiaeth wedi'u nodi isod.
- 6.2 Mae'r dystiolaeth yn awgrymu nad yw moeseg data yn bwnc statig. Bydd yn parhau i esblygu wrth i fwy o sefydliadau sefydlu eu prosesau moesegol cyd-destun penodol eu hunain, ac wrth i dechnoleg a chymdeithas newid ac wrth i ymwybyddiaeth y cyhoedd gynyddu. Felly mae angen i ddadansoddwyr ac ymchwilwyr werthfawrogiad nid yn unig o'r fframweithiau moesegol sy'n bodoli, ond hefyd gwerthfawrogiad ehangach ac atblygol o'r themâu a drafodwyd uchod, gan gynnwys effeithiau cadarnhaol a negyddol eu gwaith. Mae hyn yn cynnwys cyd-destun budd y cyhoedd y mae eu hymchwil yn eistedd ynddo ac effeithiau cymdeithasol ehangach eu gwaith (Sefydliad Ada Lovelace 2022). Mae hefyd yn cynnwys yr anghydbwysedd pŵer sy'n bodoli sy'n ei gwneud hi'n anodd i unigolion olrhain seilweithiau data mawr, neu oruchwylio ansawdd, arwyddocâd gwyddonol ac effaith gymdeithasol bosibl y wybodaeth a gynhyrchir (Leonelli 2020).
- 6.3 Mae defnydd moesegol o ddata lefel unigol yn gofyn am fwy na chydymffurfio â'r gyfraith, oherwydd nid yw hyn o reidrwydd yn sicrhau bod data'n cael eu defnyddio mewn ffyrdd sy'n dderbyniol yn gymdeithasol (Jones a Ford, 2018; Kalkman et al. 2019), nac yn sicrhau bod egwyddorion moesegol ehangach yn cael eu bodloni. Mae llenyddiaeth ar foeseg data gweinyddol yn nodi materion moesegol sy'n ymwneud â diogelwch a phreifatrwydd ond mae hefyd yn nodi pryderon eraill ynghylch diffyg ymwybyddiaeth y cyhoedd, a'r angen am ymgysylltu parhaus â'r cyhoedd ac atebolrwydd cymdeithasol (OSR 2023, Nuffield Council on Bioethics 2015). Mae'r dystiolaeth yn awgrymu bod awydd cyhoeddus am ragor o wybodaeth ynghylch arferion rhannu data a chysylltu data (Aitkin et al. 2016) a bod angen i ymchwilwyr fod yn rhagweithiol ac wedi'u

huwchsgilio wrth gyfathrebu ac ymgysylltu â chyhoeddwyr amrywiol er mwyn galluogi dealltwriaeth wirioneddol (Kennedy et al. 2022).

7. Argymhellion o'r llenyddiaeth

- 7.1 Nod yr adolygiad llenyddiaeth hwn oedd gwneud argymhellion ar arfer gorau yn seiliedig ar y dystiolaeth. Mae'r adran hon yn nodi detholiad bach o argymhellion allweddol o'r llenyddiaeth a fyddai'n ddefnyddiol i berchnogion data ac ymchwilwyr eu hystyried wrth ymgymryd ag ymchwil data gweinyddol.

Llywodraethu a Gwiriadau Atebolrwydd

- 7.2 Mae Harkness et al. yn argymhell y dylid parhau i adolygu prosesau diogelwch a llywodraethu er mwyn sicrhau eu bod yn parhau i fod yn addas at y diben wrth i dechnoleg a chanfyddiadau cymdeithasol newid a datblygu (Harkness et al. 2022). Un risg sy'n gysylltiedig â chysylltu gwahanol setiau data yw'r risg o ail-adnabod, felly mae'n rhaid i statws setiau data mawr gael ei adolygu bob amser (Laurie a Stevens 2016). At hynny, dylai pwyllgorau moeseg a chyrrff monitro annibynnol fod yn eu lle i adolygu a chymeradwyo ceisiadau mynediad at ddata (Kalkman et al. 2019).

Budd Cyhoeddus

- 7.3 Mae Scott et al. yn argymhell bod awdurdodau cyhoeddus yn mynegi ac yn cyfathrebu dimensiynau gwahanol a chynnil budd y cyhoedd mewn prosiectau ymchwil data unigol mewn ffyrdd sy'n benodol ac yn hawdd eu deall (Scott et al. 2018). Mae hyn hefyd yn gofyn am asesiad o'r risgiau a'r effeithiau negyddol er mwyn caniatáu ar gyfer ymarfer cydbwysu i asesu dimensiynau moesegol prosiectau (Y Sefydliad Data Agored).

Codi Ymwybyddiaeth a Thryloywder

- 7.4 Mae ADR UK yn argymhell bod angen cynyddu ymwybyddiaeth y cyhoedd am arferion presennol a'r defnydd o ddata (ADR UK 2020). Mae Harkness et al. (2022) yn argymhell y dylid ymarfer 'tryloywder rhagweithiol' ac y dylai ymchwilwyr fynd ymhellach na darparu gwybodaeth i'r rhai sy'n chwilio amdani. Dylid dod â gwybodaeth glir i fywydau pobl yn rhagweithiol trwy amrywiaeth o sianeli a dulliau i godi ymwybyddiaeth gyffredinol o arferion ymchwil data (Harkness et al. 2022). Mae Kennedy et al. (2022) yn amlygu bod angen cyfathrebu i fynd i'r afael â phryderon penodol sydd gan bobl am yr hyn sy'n bwysig iddynt (fel yr hyn a allai fynd o'i le, beth allai newid yn y dyfodol a sut y

gallai pobl o grwpiau lleiafrifol neu ddifreintiedig gael eu heffeithio) (Kennedy et al. 2022).

Dulliau Eang a Lluosog o Ymgysylltu â'r Cyhoedd

- 7.5 Mae Sefydliad Ada Lovelace (2022) yn eirioli dros ddefnyddio ystod eang o ddulliau ymgysylltu, cyfranogi a chyfathrebu sy'n cynnwys rheithgorau dinasyddion, paneli cyhoeddus a dulliau allgymorth i gynnwys ystod amrywiol o bobl a fydd yn helpu i hysbysu llunwyr polisi ac ymchwilwyr am yr hyn y mae pobl yn ei ddisgwyl gan ddata, ymchwil a'r defnydd ohonynt (Sefydliad Ada Lovelace 2022). Mae Aitkin et al. yn argymhell y gellid defnyddio dulliau ymchwil ansoddol fel ffurfiau o ymgysylltu i annog ymddiriedaeth y cyhoedd. Mae Harkness et al. (2022) yn argymhell bod ymchwilwyr yn cael eu hyfforddi mewn sgiliau sy'n eu galluogi i ymgysylltu â'r cyhoedd mewn ffordd ystyrlon.

Ymchwil Pellach

- 7.6 Mae angen ymchwil pellach i ddeall sut y gallai 'stiwardiaeth data cyfranogol' (Sefydliad Ada Lovelace 2021) weithio'n ymarferol mewn ymchwil data gweinyddol i archwilio cysyniadau allweddol ynghylch tegwch, cydraddoldeb a chynhwysiant.
- 7.7 Mae Sefydliad Ada Lovelace yn cyfeirio at fentrau data sy'n 'adeiladu cyhoeddusrwydd' ac sy'n rhoi digon o ddealltwriaeth i bobl ryngweithio â strwythurau data neu gymryd camau os dymunant (Sefydliad Ada Lovelace 2021). Gallai ymchwil pellach archwilio'r ffyrdd y gellid cynnwys 'cyhoeddusrwydd' mewn ymchwil data gweinyddol.

Cyfeiriadau

Ada Lovelace Institute (2021) [Participatory data stewardship \(Ada Lovelace Institute\) pdf](#). London. [Accessed: February 2025].

Ada Lovelace Institute (2022). [Looking before we leap \(Ada Lovelace Institute\)](#) London. [Accessed: February 2025].

Ada Lovelace Institute (2022) [Rethinking data and rebalancing digital power \(Ada Lovelace Institute\)](#). London. [Accessed: February 2025].

Ada Lovelace Institute (2025) [About Ada Lovelace Institute](#). London. [Accessed: February 2025].

ADR UK (2020) [Trust, security and public interest: Striking the balance \(ADR UK\)](#). Swindon. [Accessed: March 2025].

ADR UK (2025) Website [ADR UK](#). Swindon. [Accessed: February 2025].

ADR Wales (2022) [ADR Wales Programme of Work 2022 to 2026 pdf](#). Swansea and Cardiff. [Accessed: March 2025].

Aitken, M., de St. Jorre, J., Pagliari, C., Jepson, R. and Cunningham-Burley, S. (2016) [Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies \(BMC Medical Ethics\)](#). BMC Med Ethics 17:73. [Accessed: February 2025].

Britain Thinks (2021) [Complete transparency, complete simplicity: How can the public sector be meaningfully transparent about algorithmic decision making? \(Gov.UK\)](#) Thinks Insight and Strategy, London. [Accessed: February 2025].

Carter, P., Laurie, G.T. and Dixon-Woods, M. (2015). [The Social Licence for research: why care.data ran into trouble](#) Journal of Medical Ethics, 41:5 [Accessed: February 2025].

Central Digital and Data Office (2020) [Data Ethics Framework \(Gov.UK\)](#) Department for Science, Innovation and Technology, London. [Accessed: February 2025].

Centre for Data Ethics and Innovation (2020) [Addressing trust in public sector data use \(Gov.UK\)](#). Department for Science, Innovation and Technology, London. [Accessed: February 2025].

Department for Science, Innovation and Technology and Department for Digital, Culture, Media & Sport (2020) [Data Ethics and AI Landscape \(Gov.UK\)](#). London. [Accessed: February 2025].

Digital Economy Act 2017 [Research Code of Practice and Accreditation Criteria Guidance \(Gov.UK\)](#) [Accessed: February 2025].

Floridi, L. and Taddeo, M. (2016) [What is data ethics? \(Royal Society\)](#) Philosophical Transactions of The Royal Society. [Accessed: February 2025].

Government Social Research (2021) [GSR Professional Guidance Ethical Assurance for Social and Behavioural Research in Government pdf](#) HM Treasury London. [Accessed: February 2025].

Harkness, F., Blodgett, J., Rijneveld, C., Waind, E., Amugi, M. and McDonanld, F. (2022) [Building a trustworthy national data research infrastructure: A UK-wide public dialogue](#) Zenodo, Geneva. [Accessed: February 2025].

HM Treasury (2015) [Aqua Book \(Gov.UK\)](#) London. [Accessed: February 2025].

HM Treasury (2020) [Magenta Book \(Gov.UK\)](#) London. [Accessed: February 2025].

Information Commissioner's Office (2025) [Principle \(a\): Lawfulness, fairness and transparency](#) ICO. Wilmslow. [Accessed: February 2025].

Jones, K.H. and Ford, D.V. (2018) [Population data science: advancing the safe use of population data for public benefit \(Epidemiology and Health Journal\)](#) Epidemiol Health, 2018;40. [Accessed: February 2025].

Kalkman, S., Mostert, M., Gerlinger, C., van Delden, J. and van Thiel, G. (2019) [Responsible data sharing in international health research: A systematic review of principles and norms \(BMC Medical Ethics Journal\)](#) BMC Med Ethics 20, 21 [Accessed: February 2025].

Kennedy, H., Oman, S., Taylor M., Bates J. and Steedman R. (2020) [Public understanding and perceptions of data practices: A review of existing research](#) Living with Data, University of Sheffield. [Accessed February 2025].

Kennedy, H., Oman, S., Ditchfield, H., Taylor, M., Bates, J., Medina-Perea, I., Pinney, L. and Fraczak, M. (2022) [Data matters are human matters: Final Living with Data report on public perceptions of public sector data uses](#) Living with Data. University of Sheffield. [Accessed: February 2025].

Laurie G. and Stevens L. (2016) Developing a public interest mandate for the governance and use of administrative data in the United Kingdom. Journal of Law and Society. 43, 3.

Leonelli, S. (2019) [Data governance is key to interpretation: Reconceptualizing data in data science](#). Harvard Data Science Review, 1.1. [Accessed: February 2025].

Leonelli, S. (2020) [Scientific research and big data \(Stanford Encyclopaedia of Philosophy Archive\)](#) Summer 2020 edition. [Accessed: February 2025].

Muller, S., Kalkman, S., van Thiel, G., Mostert, M and van Delden, J. (2021) [The social licence for data-intensive health research: Towards co-creation, public value and trust \(BMC Medical Ethics\)](#) BMC Med Ethics 22, 110 [Accessed: February 2025].

National Data Guardian (2022) [What do we mean by public benefit? Evaluating public benefit when health and adult social care data is used for purposes beyond individual care \(Gov.UK\)](#) Office of the National Data Guardian, Leeds. [Accessed: February 2025].

Nuffield Council on Bioethics (2015) [The collection, linking and use of data in biomedical research and health care: Ethical issues](#). Nuffield Council on Bioethics. London. [Accessed: February 2025].

Office for National Statistics (2023) [What we know from engaging with the public on data \(Office for National Statistics\)](#). ONS. Newport. [Accessed: February 2025].

Office for Statistics Regulation and ADR UK (2022) [A UK-wide public dialogue exploring what the public perceive as public good use of data for research and statistics \(Office for Statistics Regulation\)](#). OSR. London. [Accessed: February 2025].

Office for Statistics Regulation (2023) [Data sharing and linkage for the public good \(Office for Statistics Regulation\) pdf](#). OSR. London. [Accessed: February 2025].

Open Data Institute (2021) [The data ethics canvas](#). ODI. London. [Accessed: February 2025].

Open Data Institute (2025) [Data ethics essentials](#). ODI. London. [Accessed: February 2025].

Organisation for Economic Cooperation and Development (2019) [Good practice principles for data ethics in the public sector pdf](#). OECD. Paris. [Accessed: February 2025].

Organisation for Economic Cooperation and Development (2019) [The path to becoming a data-driven public sector](#). OECD. Paris. [Accessed: February 2025].

Public Engagement in Data Research Initiative (2025). [PEDRI website](#). London. [Accessed: February 2025].

Royal Society (2020) [UK data governance landscape \(Royal Society\) pdf](#). London. [Accessed: February 2025].

Royal Society (2020) [The information environment and its effects on individuals and groups interdisciplinary literature review \(Royal Society\) pdf](#). London. [Accessed: February 2025].

SAIL Databank (2025) [Website](#). Swansea University, Swansea. [Accessed: February 2025].

Scott, K. (2018) [Data for public benefit: Balancing the risks and benefits of data sharing \(Involve.Org\)](#) Involve & Understanding Patient Data, London, and Carnegie UK Trust, Dunfermline [Accessed: February 2025].

Sexton, A., Shepherd, E., Duke-Williams, O. and Everleigh, A. (2017) [A balance of trust in the use of government administrative data \(Springer Journals\)](#) Arch Sci 17, 305 to 30. [Accessed: February 2025].

Stiles, P., Boothroyd, R., Robst, J. and Ray, V. J. (2011) [Ethically using administrative data in research \(Sage Journals\)](#) Administration & Society 43(2):171 to 192 [Accessed: February 2025].

UK Research and Integrity Office (2023) [Code of practice for research](#) UKRIO. London. [Accessed: February 2025].

UK Statistics Authority (2022) [Ethics self assessment tool](#) . UKSA. London. [Accessed: February 2025].

UK Statistics Authority (2022) [Ethical principles](#) .UKSA. London. [Accessed: February 2025].

UK Statistics Authority Centre for Applied Data Ethics (2021) [Considering public good in research and statistics: Ethics guidance](#) . UKSA. London [Accessed: February 2025].

Understanding Patient Data & Ada Lovelace Institute (2020). [Foundations of fairness - Where next for NHS health data partnerships? Pdf](#) Understanding Patient Data. London. [Accessed: February 2025].

Welsh Government (2021) [Digital Strategy for Wales \(Welsh Government\)](#) Cardiff. [Accessed: February 2025].

Yates, E., Addison, K. and Weedall, A. (2025) A pilot study of ethics in administrative data research: workshop analysis. Welsh Government. Available at: [Ethics in administrative data research \(Welsh Government\)](#) [Available: 30 April 2025].

Atodiad A

Organised alphabetically by source.

[Ada Lovelace Institute](#)

An independent research institute with a mission to ensure data and AI work for people and society.

Search Strategies

Browsed Library of publications. 45 results, 2 useful

[ADR UK](#)

Good government policy needs good evidence. But good quality evidence is expensive and time-consuming to produce, and evidence from academic research can also have a different focus from what government needs most to make policy decisions and deliver effective services. ADR UK's mission is to solve both these problems by transforming the way researchers access the UK's wealth of public sector data.

Search Strategies

Browsed publications. No new relevant publications (other than those provided by search requester).

[Analysis & Policy Observatory \(APO\)](#)

APO is an open access evidence platform – making public policy research and resources accessible and useable.

Search Strategies

1. "administrative data" AND ethic*. 4 results, none useful
2. data AND link* AND ethic* .19 results, none useful
3. data AND ethic*, subject filter: ethics . 21 results none useful

[Applied Social Sciences Index and Abstracts \(ASSIA\)](#)

ASSIA is an indexing and abstracting tool covering health, social services, psychology, sociology, economics, politics, race relations and education. Updated monthly, ASSIA provides a comprehensive source of social science and health information for the practical and academic professional.

Search Strategies

```
1. noft("data stewardship" OR "data ethic*" OR "data privacy" OR "data security" OR "data standard*" OR "data linking" OR "data integrity" OR "administrative data" OR "data governance") OR (MAINSUBJECT.EXACT("Data collection") OR MAINSUBJECT.EXACT("Sociodemographic data") OR MAINSUBJECT.EXACT("Datasets") OR MAINSUBJECT.EXACT("Data")) AND noft(ethic*) OR (MAINSUBJECT.EXACT("Ethical aspects") OR MAINSUBJECT.EXACT("Ethics")) AND noft(reuse OR research) OR (MAINSUBJECT.EXACT("Reuse") OR MAINSUBJECT.EXACT("Research")) AND noft("public good" OR "public benefit" OR "public interest") OR MAINSUBJECT.EXACT("Public interest") Date after: 01/01/2010
```

Notes - 9 results, one useful

```
2. title("data stewardship" OR "data ethic*" OR "data privacy" OR "data security" OR "data standard*" OR "data linking" OR "data integrity" OR "administrative data" OR "data governance") OR (MAINSUBJECT.EXACT("Data collection") OR MAINSUBJECT.EXACT("Sociodemographic data") OR MAINSUBJECT.EXACT("Datasets") OR MAINSUBJECT.EXACT("Data")) AND title(ethic*) OR (MAINSUBJECT.EXACT("Ethical aspects") OR MAINSUBJECT.EXACT("Ethics")) OR abstract(ethic*) AND noft(reuse OR research) OR (MAINSUBJECT.EXACT("Reuse") OR MAINSUBJECT.EXACT("Research")) Date after: 01/01/2010
```

Notes - 271 results, one useful

```
3. noft("public sector data" OR "administrative data" OR "government data") AND noft(ethic*) OR (MAINSUBJECT.EXACT("Ethical aspects") OR MAINSUBJECT.EXACT("Ethics")) Date after: 01/01/2010
```

Notes - 17 results, none useful

4. noft(data NEAR/2 link*) AND (noft(ethic*) AND (MAINSUBJECT.EXACT("Ethical aspects") OR MAINSUBJECT.EXACT("Ethics")))

Date after: 01/01/2010

Notes - 11 results.

DARE UK

DARE UK aims to design and deliver a coordinated and trustworthy national data research infrastructure to support cross-domain research for public good.

Search Strategies

Browsed publications - 1 useful.

Future of Privacy Forum

Future of Privacy Forum and the FPF Education and Innovation Foundation are non-profit organizations that serve as catalysts for privacy leadership and scholarship, advancing principled data practices in support of emerging technologies.

Search Strategies

Browsed white papers, filtered by issues: ethics. 5 results, 1 useful.

Google

Targeted domain searches.

Search Strategies

1. site: org.uk "administrative data" AND ethic*. Browsed first 10 pages, some useful results.

2. site: gov.uk "administrative data" AND ethic*. Browsed first 9 pages, some useful results.

3. site: ac.uk "administrative data" AND ethic*. Browsed first 10 pages, few useful results.

Organisation for Economic Cooperation and Development (OECD) iLibrary

This is the online library of the OECD featuring its books, papers and statistics and is the gateway to OECD's analysis and data.

Search Strategies

(Abstract 'data*') (Language 'en') AND (Abstract 'ethic*') AND (") published between 2010 and 2023.

22 results, 3 useful.

Open Data Institute (ODI)

The ODI is a non-profit with a mission to work with companies and governments to build an open, trustworthy data ecosystem.

Search Strategies

Browsed reports - Some useful results

ScienceDirect

ScienceDirect offers access to references, abstracts and (selected) full text for over 2,000 journals in the areas of science, technology and medicine. Our subscription includes access to the full text of all the journals within the Social Sciences Subject Collection - over 160 titles covering health, education, environmental and social justice issues.

Search Strategies

1. Title, abstract, keywords: (data) AND ethic AND (research OR reuse) Year: 2010 to 2023. 71 results, none useful.

2. Title: data AND ethics Year: 2010 to 2023. 44 results, one useful

Scopus

Scopus uniquely combines a comprehensive, expertly curated abstract and citation database with enriched data and linked scholarly literature across a wide variety of disciplines.

Search Strategies

1. (TITLE-ABS-KEY ("data steward*" OR "data ethic*" OR "data privacy" OR "data security" OR "data standard*" OR "data linking" OR "data integrity" OR "administrative data" OR "data governance") AND TITLE-ABS-KEY (ethic*) AND TITLE-ABS-KEY (reuse OR research) AND TITLE-ABS-KEY ("public good" OR "public benefit" OR "public interest")) AND PUBYEAR > 2009 AND PUBYEAR > 2009.

Notes

27 results, 2 useful

2. (TITLE-ABS-KEY ("data linking" OR "administrative data")) AND TITLE-ABS-KEY (ethic*) AND TITLE-ABS-KEY (reuse OR research)) AND PUBYEAR > 2009

Notes

157 results, several useful

3. TITLE-ABS-KEY (= ("public sector data" OR "government data")) AND TITLE-ABS-KEY (ethic*) AND PUBYEAR > 2009

Notes

62 results, none useful

4. TITLE-ABS-KEY (= (data W/2 link*)) AND TITLE-ABS-KEY (= (ethic*)) AND TITLE-ABS-KEY (= ("public good" OR "public benefit" OR "public interest")) AND PUBYEAR > 2009

Notes

19 results, 1 useful

[The Alan Turing Institute](#)

The UK's national institute for data science and artificial intelligence.

Search Strategies

Browsed publications filtered by research area: ethics. 23 results, 1 useful

[UK Data Service](#)

Trusted access and training to use the UK's largest collection of economic, population and social research data for teaching, learning and public benefit.

Search Strategies

Browsed current and past projects. No useful results.

[Web of Science](#)

Web of Science is multidisciplinary and international in scope. It indexes over 21,000 journals across 4 key databases: Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index and the Emerging Sources Citation Index.

Search Strategies

1. ((TS=("data stewardship" OR "data ethic*" OR "data privacy" OR "data security" OR "data standard*" OR "data linking" OR "data integrity" OR "administrative data" OR "data governance") OR TI=("data stewardship" OR "data ethic*" OR "data privacy" OR "data security" OR "data standard*" OR "data linking" OR "data integrity" OR "administrative data" OR "data governance") OR AB=("data stewardship" OR "data ethic*" OR "data privacy" OR "data security" OR "data standard*" OR "data linking" OR "data integrity" OR "administrative data" OR "data governance")) AND (TS=(ethic*) OR TI=(ethic*) OR AB=(ethic*)) AND (TS=(reuse OR research) OR TI=(reuse OR research) OR AB=(reuse OR research)) AND (TS=("public good" OR "public benefit" OR "public interest") OR TI=("public good" OR "public benefit" OR "public interest") OR AB=("public good" OR "public benefit" OR "public interest"))))

Notes

15 results, 1 useful

2. ((TS=(data) OR AB=(data) OR TI=(data)) AND (TS=(ethic*) OR TI=(ethic*) OR AB=(ethic*)) AND (TS=(reuse OR research) OR TI=(reuse OR research) OR AB=(reuse OR research)) AND (TS=("public good" OR "public benefit" OR "public interest") OR TI=("public good" OR "public benefit" OR "public interest") OR AB=("public good" OR "public benefit" OR "public interest"))))

Date: 01/01/2010 – present

Notes

107 results, 3 useful

3. (TS=("public sector data" OR "administrative data" OR "government data") OR TI=("public sector data" OR "administrative data" OR "government data") OR AB=("public sector data" OR "administrative data" OR "government data")) AND (TS=(ethic*) OR TI=(ethic*) OR AB=(ethic*))

Date: 01/01/2010 – present

Notes

184 results, 4 useful

4. (TS=(data NEAR/2 link*) OR TI=(data NEAR/2 link*) OR AB=(data NEAR/2 link*)) AND (AB=(ethic*) OR TI=(ethic*) OR TS=(ethic*)) AND (TS=("public good" OR "public benefit" OR "public interest") OR AB=("public good" OR "public benefit" OR "public interest") OR TI=("public good" OR "public benefit" OR "public interest"))

Date: 01/01/2010 - present

Notes

13 results, 2 useful

Atodiad B

Priority 1, 2, 3.	Title	Author	Citation and Link	Year of Publ.	Publ. type
0.1	Ethical Principles	UK Statistics Authority	UK Statistics Authority	2023	Guidance
0.1	A UK-wide public dialogue exploring what the public perceive as 'public good' use of data for research and statistics	Harkness, F, Rijnveld, C, Liu, Y et al	Kohlrabi Consulting, Office for Statistics Regulation , ADRUK, Economic and Social Research Council	2022	Report
0.1	What is research integrity?	The University of Edinburgh	The University of Edinburgh	2021	Guidance
0.1	Public dialogue on location data ethics: Engagement report	Maxwell, M, McCool, S Peppin, A et al	Traverse, Ada Lovelace Institute, Geospatial Commission, Gov.Uk	2021	Report
0.1	Trust, Security and Public Interest: Striking the Balance: A review of previous literature on public attitudes towards the sharing and linking of administrative data for research	ADRUK	ADR UK (Administrative Data Research UK) Strategic Hub Economic & Social Research Council	2020	Review article
0.1	Data Ethics Framework	Centra Digital and Data Office	Gov.Uk	2018	Guidance
0.1	Ethically Using Administrative Data in Research: Medicaid Administrators' Current Practices and Best Practice Recommendations	Stiles, P, Boothroyd, R, Robst, J and V. Ray, J	Administration & Society 43(2), pages 171 to 192 (Sage Journals)	2011	Journal article
0.1	An EU code of ethics for socio-economic research	Dench, S, Iphofen, R, Huws, U	Institute for Employment Studies, Report 412	2004	Report
1	Attitudes Toward Data Linkage: Privacy, Ethics, and the Potential for Harm	Fobia, A, Childs, J Eggleston, C	Craig A. Hill, Paul P. Biemer, et al. (ed.) Big Data Meets Survey Science: A Collection of Innovative Methods pages 683 to 712 (Wiley Online Library)	2020	Book chapter
1	Addressing trust in public sector data use	Centre for Data Ethics and Innovation	Gov.Uk	2020	Report
1	The role of data in building trust	OECD	In The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies.	2019	Book chapter

Priority 1, 2, 3.	Title	Author	Citation and Link	Year of Publ.	Publ. type
1	The role and nature of consent in government administrative data	Sexton, A, Shepherd, E, Duke-Williams, O and Eveleigh, A	Big Data & Society 5(2)	2018	Journal article
1	Privacy Protective Research: Facilitating Ethically Responsible Access to Administrative Data	Goroff, D, Polonetsky, J and Tene, O	The ANNALS of the American Academy of Political and Social Science 675(1), pages 46 to 66	2017	Journal article
1	A balance of trust in the use of government administrative data	Sexton, A, Shepherd, E, Duke-Williams, O and Eveleigh, A	Archival Science 17, pages 305 to 330	2017	Journal article
1	Developing a Public Interest Mandate for the Governance and Use of Administrative Data in the United Kingdom	Laurie G and Stevens, L	Journal of Law and Society 43(3), pages 360 to 392	2016	Journal article
1	Ethical Use of Administrative Data for Research Purposes	Stiles, P and Boothroyd, R	From Actionable Intelligence: using integrated data systems to achieve a more effective, efficient and ethical government	2015	Book chapter
2	Building a trustworthy national data research infrastructure: A UK-wide public dialogue	Harkness, F, Jo Blodgett, J, Rijnveld, C et al	Zenodo	2022	Report
2	The social licence for data-intensive health research: towards co-creation, public value and trust	Laurie, S, Kalkman, S, van Thiel, G et al	BMC Medical Ethics 22, article number:110	2021	Journal article
2	The Ethics of Consent in a Shifting Genomic Ecosystem	Soo-Jin Lee, S	Annual Review of Biomedical Data Science 4, pages 145 to 164	2021	Review article
2	The evolution of data ethics in clinical research and drug development	Heidelberg, M, Kelman, A, Hopkins, J, Allen, M.E	Ethics, Medicine and Public Health 14, 100517	2020	Journal article
2	The role of data custodians in establishing and maintaining social licence for health research	Allen, J, Adams, C, Flack, F	Bioethics 33(4), pages 502 to 510	2019	Journal article
2	Responsible data sharing in international health research: a systematic review of principles and norms	Kalkman, S, Mostert, M, Gerlinger, C et al	BMC Medical Ethics 20, article number: 21 (open access)	2019	Review article
2	Population data science: advancing	Jones, K, Ford, D	Epidemiology and Health 40, e2018061	2018	Review article

Priority 1, 2, 3.	Title	Author	Citation and Link	Year of Publ.	Publ. type
	the safe use of population data for public benefit				
2	The opportunities and ethics of big data: practical priorities for a national Council of Data Ethics	Varley-Winter, O and Shah, H	Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374(2083)	2016	Journal article
2	Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies	Aitken, M, De St Jorre, J , Pagliari, C et al	BMC Medical Ethics 17(1), pages 1 to 24, article number: 73	2016	Review article
3	Participatory data stewardship: A framework for involving people in the use of data	Ada Lovelace Institute	Ada Lovelace Institute	2021	Report
3	Design for data ethics: using service design approaches to operationalize ethical principles on four projects	Drew, C	Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 276(2128)	2018	Journal article
3	Data for Public Benefit: Balancing the risks and benefits of data sharing	Scott, K et al	Involve, Understanding Patient Data, Carnegie UK Trust	2018	Report
3	Challenges in administrative data linkage for research	Harron, K, Dibben, C, Boyd, J Hjern, A, et al	Big Data & Society 4(2)	2017	Journal article

The Welsh Government Library provides an internal information service to Welsh Government Ministers and officials. We also maintain the publications archive available through the publications catalogue. Our remit means that we are not able to provide additional library or information services to external stakeholders or members of the public. This includes help accessing references that are not available in full text within these search results. This report is not exhaustive and inclusion in the service is not an endorsement of an organisation's website or research.

Atodiad C

Ada Lovelace Institute

Looking before we leap: Expanding ethical review processes for AI and data science research (2022).

Ada Lovelace Institute

Rethinking data and rebalancing digital power (2022).

Britain Thinks

Complete transparency, complete simplicity: how can the public sector be meaningfully transparent about algorithmic decision making? (2021).

Central Digital and Data Office

Understanding artificial intelligence ethics and safety (2019).

Data Ethics EU

EU data ethics impact assessment (2021).

Department for Science, Innovation and Technology and Department for Digital, Culture, Media & Sport

Data ethics and AI guidance landscape (2020).

Floridi, L and Taddeo, M

What is data ethics? (2021).

Government Social Research

GSR professional guidance ethical assurance for social and behavioural research in government (2021).

HM Treasury

Aqua Book: guidance on producing quality analysis for government (2015).

HM Treasury

Magenta Book: central Government guidance on evaluation (2020).

Kennedy, H et al

Public understanding and perceptions of data practices: a review of existing research (2020).

Kennedy, H et al

Data matters are human matters: final living with data report on public perceptions of public sector data uses (2022).

Leonelli, S

Data: from objects to assets (2019).

Leonelli, S

Data governance is key to interpretation: reconceptualizing data in data science (2019).

National Data Guardian

What do we mean by public benefit? (2022)

Nuffield Council on Bioethics

The collection, linking and use of data in biomedical research and health care: ethical issues (2015).

OECD

Good practice principles for data ethics in the public sector (2019).

Office for National Statistics

What we know from engaging with the public on data: June 2023

Office for Statistics Regulation

Data sharing and linkage for the public good (2023).

Open Data Institute

The Data Ethics Canvas.

Royal Society

The UK data governance landscape (2020).

Royal Society

The information environment and its effects on individuals and groups: literature review (2020).

UK Research and Integrity Office

Code of Practice for Research (2023).

UKRI

Framework for Research Ethics (2021).

Understanding Patient Data and Ada Lovelace Institute

Foundations of fairness - Where next for NHS health data partnerships? (2020)

UK Statistics Authority

Ethics Self-Assessment Tool (2022).

UK Statistics Authority

Ethical Principles (2022).

UK Statistics Authority Centre for Applied Data Ethics

Considering public good in research and statistics: ethics guidance (2021).

UK Statistics Authority Centre for Applied Data Ethics

Considering public views and engagement regarding the use of data for research and statistics (2021).