PRIFYSGOL ABERYSTWYTH UNIVERSITY

SEMESTER UN 2025

CYFRIFIADUREG

CC27020 : MODELU DATA PARHAUS

Amser a ganiateir: **2 AWR**

Atebwch bob un o'r **TRI CHWESTIWN**

Mae pob cwestiwn efo'r un marc.

CEWCH DDOD Â'R CANLYNOL I MEWN I'R ARHOLIAD:
HYD AT 3 TUDALEN O NODIADAU AR BABUR A4 (A ALL FOD YN DDWYOCHROG). NID OES UNRHYW GYFYNGIADAU YCHWANEGOL AR GYNNWYS NA FFORMAT Y NODIADAU HYN. DIM DYFEISIAU TRYDANOL.

NI CHANIATEIR CYFRIFANELLAU.

MAE MODD I FYFYRWYR GYFLWYNO ATEBION I'R PAPUR HWN NAILL AI YN Y GYMRAEG NEU'R SAESNEG

Wedi eistedd, gall myfyrwyr gwblhau clawr ffrynt y llyfr(au) atebion a'r ffurflen presenoldeb YN UNIG.

PEIDIWCH A THROI'R PAPUR CWESTIYNNAU HYD NES RHODDIR CYFARWYDDYD.

1. **Mae'r cwestiwn hyn yn ymneud â cywirdeb data ac SQL.**

a) Mae gwasanaeth benthyca gliniaduron yn storio gwybodaeth am fenthyca gliniaduron. Mae'r tabl Gliniaduron yn dal manylion dyfeisiau. Mae'r tabl Benthyciadau yn cofnodi pob benthyciad. Isod mae'r diffiniad SQL sgerbwd a roddir ar gyfer perthnasoedd Gliniaduron a Benthyciadau.

```
CREATE TABLE laptops (
  laptop_id INT,
  model VARCHAR(100)
);

CREATE TABLE loans (
  loan_id INT,
  laptop_id INT,
  user_id VARCHAR(10),
  loan_date DATE,
  due_date DATE,
  return_date DATE
);
```

(i) Ychwanegwch gyfyngiadau sylfaenol gan ddefnyddio SQL DDL yn unig (dim swyddogaethau na sbardunau) i orfodi'r rheolau uniondeb isod. Gweithredwch y canlynol (ni fydd gwyriadau bach o gystrawen SQL yn cael eu cosbi):

- Allweddi cynradd ar y ddau dabl.
- Rhaid i Loans.laptop_id gyfeirio at Laptops.laptop_id.
- Yn y tabl Benthyciadau (Loans table), NID yw laptop_id, user_id, loan_date, due_date yn NULL.
- Yn y tabl Benthyciadau (Loans table), rhaid i'r due_date fod ar/ar ôl loan_date, a rhaid i'r return_date fod yn NULL (ar gyfer benthyciadau gweithredol) neu ar/ar ôl loan_date.

[9 marc]

(ii) Argymhellwch un ffordd i sicrhau uchafswm o un benthyciad gweithredol fesul gliniadur (`return_date IS NULL`). Disgrifiwch y dull yn fyr a chyfiawnhewch eich dewis. Nid oes angen cod gweithredu. [3 marc]

b) Mae rheoli trafodion a rheoli cydamseroldeb yn swyddogaethau RDBMS allweddol ar gyfer sicrhau cywirdeb data.

(i) Ystyriwch yr amserlen rhyngblethedig ganlynol S ar gyfer trafodion T_1 a T_2.

$$S = [W_1(B),\ R_2(B),\ R_1(A),\ W_2(B)]$$

Yma, mae $R_i(X)$ a $W_i(X)$ yn dynodi darlleniad ac ysgrifennu o eitem X trwy drafodiad T_i.

A yw'r amserlen hon yn gyfresiadwy? Cyfiawnhewch eich ateb. [6 marc]

(ii) Tybiwch fod angen i chi weithredu ymholiad un datganiad yn PostgreSQL i restru statws cyfredol pob gliniadur (gweler Cwestiwn 1a). Dewiswch lefel ynysu briodol o: READ COMMITTED, REPEATABLE READ, neu SERIALIZABLE, a chyfiawnhewch eich dewis. [3 marc]

c) Ystyriwch y tablau `Laptops` a `Loans` canlynol, sy'n glynu wrth y cyfyngiadau cywirdeb penodedig yng Nghwestiwn 1a.

Laptops Table

laptop_id	model
1	Dell X1
2	Dell X1
3	Apple MacBook
4	Lenovo ThinkPad

Loans Table

loan_id	laptop_id	user_id	loan_date	due_date	return_date
1001	1	abc1	2025-01-10	2025-04-10	2025-04-10
1002	1	abc1	2025-04-10	2025-07-10	2025-07-10
1003	1	abc1	2025-07-10	2025-10-10	NULL
1004	2	cat2	2025-03-05	2025-09-05	NULL
1006	2	mad2	2024-09-01	2025-03-01	2025-02-15
1005	4	sun1	2025-08-15	2025-11-15	2025-10-31

(i) Rhestrwch y canlyniadau ar gyfer yr ymholiad SQL canlynol, sydd â'r bwriad o ddod o hyd i bob `user_ids` sydd wedi cael model `Dell X1` ar fenthyg iddynt. [4 marc]

```
SELECT DISTINCT T2.user_id
FROM Laptops T1
INNER JOIN Loans T2 ON T1.laptop_id = T2.laptop_id
WHERE T1.model = 'Dell X1';
```

(ii) Beth yw allbwn yr ymholiad o Gwestiwn 1c(i) os caiff `INNER JOIN` ei ddisodli gan `LEFT OUTER JOIN`? [4 marc]

(iii) Ystyriwch y ddau Ymholiad SQL A a B a restrir isod, y ddau wedi'u bwriadu i ddod o hyd i ddefnyddwyr sy'n gysylltiedig â benthyciadau lluosog. Eglurwch y gwahaniaeth rhwng y canlyniadau y byddai Ymholiad A ac Ymholiad B yn eu cynhyrchu, a nodwch beth mae pob ymholiad yn ei gyfrif mewn gwirionedd. [4 marc]

Ymholiad A

```
SELECT user_id
FROM Loans
GROUP BY user_id
HAVING COUNT(loan_id) > 1;
```

Ymholiad B

```
SELECT user_id
FROM Loans
GROUP BY user_id
HAVING COUNT(DISTINCT laptop_id) > 1;
```

2. **Mae'r cwestiwn hyn yn ymneud â theori a Normaleiddio (*Normalization*).**

Mae Tabl 1 yn dangos data sampl ar ymweliadau cleifion ag adrannau amrywiol ysbyty. Mae pob rhes yn cynrychioli diagnosis neu driniaeth a roddwyd yn ystod ymweliad un claf.

Table 1: Tabl Hospital_Visits (Data Enghreifftiol)

Visit ID	Patient ID	Patient Name	VisitDate	Doctor ID	Doctor Name	Department Name	Diagnosis	Treatment Code	Treatment name
V101	PAT1	John Smith	2025-10-01	DOC1	Evans	Cardio	Hypertension	C5	ECG
V101	PAT1	John Smith	2025-10-01	DOC1	Evans	Cardio	Hypertension	C2	BP
V102	PAT2	Jane Doe	2025-10-01	DOC5	Patel	Ortho	Ankle Injury	C3	X-Ray
V103	PAT1	John Smith	2025-10-01	DOC1	Evans	Cardio	Follow-up	C2	BP
V104	PAT3	Peter Jones	2025-10-02	DOC5	Patel	Ortho	Knee Pain	C3	X-Ray
V104	PAT3	Peter Jones	2025-10-02	DOC5	Patel	Ortho	Knee Pain	C1	MRI

Rhagdybiaethau:

- Mae PatientID yn adnabod pob claf yn unigryw (PK)
- Mae DoctorID yn adnabod pob meddyg (PK) yn unigryw.
- Mae VisitID yn adnabod pob ymweliad claf (PK) yn unigryw.
- Mae TreatmentCode yn nodi pob gweithdrefn driniaeth (PK) yn unigryw.
- Mae pob meddyg yn perthyn i un adran yn union.

Dibyniaethau swyddogaethol allweddol:

- VisitID $\rightarrow$ {PatientID, PatientName, VisitDate, DoctorID, DoctorName, DepartmentName}
- PatientID $\rightarrow$ PatientName
- DoctorID $\rightarrow$ {DoctorName, DepartmentName}
- TreatmentCode $\rightarrow$ TreatmentName
- {VisitID, TreatmentCode} $\rightarrow$ Diagnosis

Nodyn: Ar gyfer pob rhan o'r cwestiwn hwn, darparwch benawdau'r sgema cronfa ddata neu'r tabl yn unig, gan nodi'n glir yr holl Allweddi Cynradd (*Primary Keys (PK)*) ac Allweddi Tramor (*Foreign Keys (FK)*). Nid oes angen llenwi'r tablau â data.

a) Eglurwch y berthynas rhwng uwch-allwedd *superkey* ac allwedd ymgeisydd *candidate key*. [3 marc]

b) Nodwch yr allwedd(au) ymgeisydd ar gyfer Tabl 1. [4 marc]

c) Gan ddefnyddio Tabl 1, nodwch ac esboniwch un math o anomaledd data a allai ddigwydd oherwydd ei strwythur. Rhowch enghraifft benodol sy'n gysylltiedig â data'r ysbyty. [4 marc]

d) Pa ffurf arferol *Normal Form* sydd ar Dabl 1 ar hyn o bryd? Cyfiawnhewch eich ateb. [4 marc]

e) A yw Tabl 1 ar yr Ail Ffurf Normal (*Second Normal Form (2NF)*)? Os na, perfformiwch y normaleiddio *normalization* cam wrth gam o Dabl 1 i 2NF. Dangoswch yn glir y sgema sy'n nodi Allweddi Cynradd *Primary Keys (PK)* ac Allweddi Tramor *Foreign Keys (FK)*, a chyfiawnhewch eich newidiadau.

f) A yw Tabl 1 ar yr Ail Ffurf Normal (*Second Normal Form (2NF)*)? Os na, perfformiwch y normaleiddio cam wrth gam o Dabl 1 i 2NF. Dangoswch yn glir y sgema sy'n nodi PKs ac FKs, a chyfiawnhewch eich newidiadau. [7 marc]

g) A yw'r tablau a gynhyrchwyd yng Nghwestiwn 2e yn y Drydedd Ffurf Normal *Third Normal Form (3NF)*? Os na, trowch y sgema *Second Normal Form (2NF)* yn *Third Normal Form (3NF)*. Cyfiawnhewch unrhyw newidiadau a wnaed. [7 marc]

h) A yw'r holl dablau yn eich sgema yn y Drydedd Ffurf Normal (*Third Normal Form (3NF)*) terfynol (o Gwestiwn 2f) hefyd ar Ffurf Boyce-Codd Normal Form (BCNF)? Eglurwch pam neu pam lai ar gyfer pob tabl. [4 marc]

3. **Mae'r cwestiwn hyn yn ymneud â Ddata Lled-Strwythuredig (*Semi-Structured Data*) a NoSQL.**

a) Archwiliwch y ffeil XML ganlynol sy'n rhestru manylebau ar gyfer y gliniaduron a ddangosir yng Nghwestiwn 1c).

```
<laptops>
  <laptop id="1" model="Dell X1" specs="">
  </laptop>
  <laptop id="2" model="Dell X1" serial="DX1-002">
    <specs>
      <ram unit="GB">32</ram>
    </specs>
  </laptop>

  <laptop id="3" brand="Apple" model="MacBook">
    <specs chipset="M2" ramGB="8"/>
  </laptop>

  <laptop id="4" model="ThinkPad" brand="Lenovo">
    <specs>
      <cpu>Ryzen 7 7840U</cpu>
      <ram unit="GB">16</ram>
      <storage type="SSD" capacityGB="512"/>
      <os>Windows 11</os>
    </specs>
  </laptop>
</laptops>
```

(i) Rhestrwch ddwy nodwedd o'r XML uchod sy'n ei wneud yn lled-strwythuredig, a rhowch enghraifft fer ar gyfer pob un. [4 marc]

(ii) Mae XML yn defnyddio elfennau (*elements*, tagiau (*tags*, a phriodoleddau (*attributes*. Cysylltwch bob eitem â'i chategori (elfen, tag, neu briodwedd).

(1) `model="Dell X1"`

(2) `<ram unit="GB">16</ram>`

(3) `<laptop>`

[3 marc]

(iii) Mae JSON yn ddewis arall yn lle XML ar gyfer data lled-strwythuredig. Ysgrifennwch gynrychiolaeth JSON ar gyfer y gliniadur gyda `id = 3`.

[5 marc]

(iv) Gan ddefnyddio'r gwrthrych JSON ar gyfer y gliniadur `id = 3` o Gwestiwn 3a(iii), llenwch y mynegiadau JSONPath i echdynnu:

```
(1) the chipset
$._______

(2) the full specs as a JSON object
$._______
```

[4 marc]

b) Mae platfform gweminar yn rhedeg ei wasanaethau (ffrwd fideo fyw a sgwrs) ar draws dau ranbarth, gydag atgynhyrchu anghydamserol. Yn ystod rhaniad rhwydwaith, mae'r ffrwd fyw yn parhau mewn un rhanbarth, ond mae'r gynulleidfa yn y llall yn colli'r diweddariadau ffrwd.

(i) Nodwch theorem y CAP a, o dan raniad, pa ddau o'r tri phriodwedd y gall y platfform eu darparu o hyd. [4 marc]

(ii) Ar gyfer y nodwedd sgwrsio byw, a ddylai'r platfform flaenoriaethu CP neu AP yn ystod y rhaniad? Rhowch gyfiawnhad byr. [4 marc]

c) Gwerthuswch y datganiadau canlynol ynghylch systemau NoSQL a'u modelau data. Nodwch a yw pob datganiad yn Wir neu'n Anghywir. Eglurwch eich atebion yn fyr. [9 marc]

S1. Nid oes angen sgema wedi'i ddiffinio ymlaen llaw ar gyfer gwerthoedd ar storfa allwedd-gwerth (*key-value store*).

S2. Mae MongoDB yn gorfodi cyfyngiadau allweddi tramor (*foreign key constraints*) rhwng casgliadau yn frodorol

S3. Mae platfform gweminar yn ymdrin â phigyn traffig 10x trwy ychwanegu 8 gweinydd ap union yr un fath gyda'r nod o gynyddu ei raddadwyedd fertigol.

DIWEDD Y PAPUR