Calculators may be used in this examination provided they are not capable of being used to store alphabetical information other than hexadecimal numbers

UNIVERSITY OF BIRMINGHAM

School of Computer Science

Storing and Managing Data

Main Summer Examinations 2024

Time allowed: 2 hours

[Answer all questions]

Turn Over

Note

Answer ALL questions. Each question will be marked out of 20. The paper will be marked out of 60, which will be rescaled to a mark out of 100. The list of GDPR articles is in the Appendix.

Question 1

(a) Given the following relational database, where primary and foreign keys are denoted via PK and FK, respectively:

patient (patientName, patientSurname, patientID PK, doctorAppointments FK, patientDateOfBirth, patientHeight, patientWeight)
appointment (appointmentCode PK, doctorID PK, doctorName, doctorSurname, roomNo)

Determine if it is in 2NF and, if not, normalise it to the 2NF and explain why/why not. For every functional dependency of the form $X \rightarrow A$, provide an example of the relation with data to show that A is partially functionally dependent on X.

Hint: recall that an attribute A is functionally dependent on X if, for every value X can take, there is one and only one corresponding value that A can take.

[10 marks]

(b) Given the following database about museums and paintings, where primary and foreign keys are denoted via PK and FK, respectively:

museum (mCode PK, mName, street, city)
room (mCode PK FK, roomNo PK, floorArea, numberOfDoors, guardian FK)
painting (pCode PK, pName, painter, dateOfPurchase, museum FK, roomNo FK)
employee (NINo PK, eName, eSurname, museum FK, startDate, role, salary)

Construct the following SQL queries:

(i) Find the highest salary of every employee role at the *Louvre* museum. **[5 marks]**

(ii) Find the average number of paintings per room in the *Momart* museum. **[5 marks]**

Question 2

Lucy is the HR manager at a software development company in France. The company is expanding its team and wants to hire new developers to work on its latest project. Lucy and her team want to ensure they hire individuals who will work well together and share similar interests and values. They consider asking applicants about various personal traits during the interview process, including technical qualifications, hobbies, favourite political party, religious beliefs, health information, sexual orientation, and ethnic origin.

Lucy wants to know which specific categories of data the company should not collect if they want to comply with the GDPR. Explain your answer using the IRAC format. **[20 marks]**

Hint: IRAC stands for Issue, Rule, Application, Conclusion.

Question 3

(a) A hash function is a powerful tool that data scientists often use in their data analytics tasks.

(i) List the names and a brief description of each of the four key properties for the hash function.

[3 marks]

(ii) List the average complexity of the hash function for *Search*, *Insert*, and *Delete*.

[3 marks]

(b) Let T be a hash-table of size 11 with the hash function $h(x) = x \bmod 11$, where x is the key ($a \bmod b$ is the remainder of the integer division $\frac{a}{b}$).

(i) Write down the entries of the hash table T after the keys 5, 19, and 15 have been inserted in that order using the following format (the first one has been completed for you):

[6 marks]

$h(x)$	0	1	2	3	4	5	6	7	8	9	10
T						5					

Hint: For example, when inserting key $x = 5$, then we have $h(x) = 5$, hence the key 5 should be placed under the position labeled as 5 (i.e., the 6th position) in T.

(ii) Now, suppose that we want to additionally insert the following keys: 27, 30, 34, 26, 12, and 21. In this case, there will be collisions. Write down the entries of the hash table T after the above have been inserted in that order using the same way as before and resolve collisions using linear probing.

[8 marks]

Appendix

This document contains the list of GDPR articles in the order they were presented in class.

- **Art. 1 - Goals of the Regulation**: Fundamental rights protection, free movement of personal data (goods and services).
- **Art. 2 - What GDPR Regulates**: "The processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system". Main exceptions: purely personal activities + law enforcement.
- **Art. 3 - Where GDPR Applies**: Data processed by processor who is in the EU, data processed of data subjects who are in the EU, even if the processor is not in the EU.
- **Art. 4 - Definitions**: personal data = "Any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".
  Processing = "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction".
- **Art. 5 - Six Principles of Data Processing**:
  (a) Lawfulness, Fairness, Transparency - Personal data shall be "processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency')". Processing of personal data is unlawful, unless based on one of six legal bases.
  (b) Purpose Limitation - Personal data shall be "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes (. . .) ('purpose limitation')".
  (c) Data Minimisation - Personal data shall be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')".
  (d) Accuracy - Personal data shall be "accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy')".

(e) Storage Limitation - Personal data shall be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; (. . .) ('storage limitation')".

(f) Integrity and Confidentiality - Personal data shall be: "processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality')".

- **Art. 6 - Lawfulness of Data Processing**: Processing of personal data is prohibited except in six situations: 1. consent, 2. necessity for contract, 3. legal obligation, 4. vital interest of a natural person, 5. necessary for performance of public task, 6. Legitimate interests of the controller.

- **Art. 7 - Consent**: Data subject's consent must be demonstrated by controller. Consent must be informed, affirmative, unambiguous, and freely given, no "take it or leave it" or threats, no opt-out. Data subject can withdraw consent at any time, this must be as easy as giving consent. **Legitimate Interests:** Most common basis for processing. Legitimate interests of the controller/third party, except where overridden by interests or fundamental rights of data subject.

- **Art. 9 - Special Categories of Data**: Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.

- **Art. 12 - Transparent Information**: The data controller shall provide easily accessible and clear information to the data subject. The data controller shall ensure that the data subject can exercise their data subject rights.

- **Art. 13-15 - Information Rights**: Right to information about: The data controller, contact details. What kind of data is being processed. The purpose and means of data processing. The legal basis for data processing. How data will be stored. Whether any automated decisions are being made. Right to copy of data.

- **Art. 16 - Right to Rectification**: Right to rectify any incorrect or incomplete information.

- **Art. 17 - Right to Erasure ("Right to be Forgotten")**: Right to have personal data erased if there is no longer a legal basis for the data to be in the hands of the processor, if data is no longer necessary for the original purpose, if consent is withdrawn, if there are no longer legitimate interests, etc. The processor needs to take reasonable steps to erase the data in all places.

- **Art. 20 - Right to Data Portability**: "The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided".
- **Art. 21 - Right to Object**: Right to object to data processing, data controller must stop processing (subject to exceptions).
- **Art. 22 - Automated Decision-Making and Profiling**: "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her". There must be a human to speak to (subject to exceptions).
- **Art. 77 - Right to Lodge a Complaint**: Data subjects have the right to lodge a complaint with the supervisory authority.
- **Art. 78 - Right to Judicial Remedy**: If the data subject disagrees with the supervisory authority's decision, they have the right to go to a judge in the Member State (processors also have this right).
- **Art. 24 - Controller Obligations**: Controller's responsibility to comply with the Regulation, adhere to the data protection principles, respect data subject rights, implement appropriate data protection policies.
- **Art. 25 - Data Protection by Design and Default**: Data controllers must, by design and default, protect the data protection principles (data minimisation, purpose limitation, etc.) and guarantee the data subject rights.
- **Art. 28 - Processors**: Data processing may be "outsourced" to a data processor, but the controller is still responsible. Must be a legal agreement in place specifying the details of the data processing.
- **Art. 30 - Records of Processing**: Both controllers and processors need to keep detailed records of their processing activities.
- **Art. 32 - Security of Processing**: Processing must be appropriately secure, e.g., pseudonymisation and encryption (technical measures) as well as organisational measures (who has access to the data?, how can they be held accountable?, etc.).
- **Art. 33-34 - Notification of Data Breach**: If data breach is likely to present a risk to fundamental rights, the breach must be communicated to the supervisory authority and the affected data subjects.

- **Art. 35 - Data Protection Impact Assessment**: If data processing is likely to form a risk to fundamental rights, the processor must perform a Data Protection Impact Assessment.
- **Art. 36 - Consultation**: If the DPIA shows that there is a high risk, the processor must consult with the supervisory authority. Large scale processing, sensitive data, important consequences.
- **Art. 51 - Supervisory Authority**: One independent supervisory authority in every Member State, Information Commissioner's Office (ICO) in the UK, responsible for monitoring compliance with the Regulation.
- **Art. 57 - Tasks of Supervisory Authority**: Monitor and enforce GDPR, carry out investigations, handle complaints from data subjects, give advice to the public, data processors, governments.
- **Art. 37 - Data Protection Officers**: Must be appointed if: public authority, large scale processing, sensitive data.

End of Paper

This page intentionally left blank.

Do not complete the attendance slip, fill in the front of the answer book or turn over the question paper until you are told to do so

Important Reminders

- Coats/outwear should be placed in the designated area.
- Unauthorised materials (e.g. notes or Tippex) must be placed in the designated area.
- Check that you do not have any unauthorised materials with you (e.g. in your pockets, pencil case).
- Mobile phones and smart watches **must** be switched off and placed in the designated area or under your desk. They must not be left on your person or in your pockets.
- You are not permitted to use a mobile phone as a clock. If you have difficulty seeing a clock, please alert an Invigilator.
- You are not permitted to have writing on your hand, arm or other body part.
- Check that you do not have writing on your hand, arm or other body part – if you do, you must inform an Invigilator immediately
- Alert an Invigilator immediately if you find any unauthorised item upon you during the examination.

Any students found with non-permitted items upon their person during the examination, or who fail to comply with Examination rules may be subject to Student Conduct procedures.