

4.1 Ethical and Social Issues Around Data Collection

Name: _____ Class: _____ Date: _____

Total: 8 marks

Objective

Build the skills to answer exam questions on **ethical and social issues around data collection**.

You must be able to:

- describe how programs collect, store, and use large amounts of **data** 数据
- explain why **privacy** 隐私 matters when personal data is gathered
- discuss the risk of **bias** 偏见 when a data set is unrepresentative

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ Data and privacy

Programs can collect, store, and process **huge** amounts of data. When that data is **personal**, **privacy** matters—it can be misused if exposed or shared without consent.

■ Bias in data

If a data set does **not represent everyone fairly**, conclusions drawn from it can be **biased**, harming under-represented groups.

2 Practice

2.1 State why privacy matters when collecting personal data. [1]

2.2 State one risk if a data set does not represent everyone fairly. [1]

2.3 State one responsibility when handling collected data. [1]

3 Exam-style questions

3.1 A data set that leaves out some groups can introduce [1]

- **A** speed
 - **B** bias
 - **C** encryption
 - **D** an array
-

3.2 Collecting personal data raises concerns about [1]

- **A** privacy
 - **B** compilation
 - **C** casting
 - **D** recursion
-

3.3 An app collects users' location data.

(a) State one privacy concern. [1]

(b) State one risk of an unrepresentative data set. [1]

(c) State one safe practice for handling the data. [1]

4 Go further

- work through the **4.1 Ethical and Social Issues Around Data Collection** lesson on the **Learn** page;
- read the **Array** section of the AP Computer Science A handout on the **Know** page.

Solutions

2.1 personal data can be misused if it is exposed or shared without consent.

2.2 the conclusions can be biased and unfair to under-represented groups.

2.3 protect privacy, store data securely, or use representative data (any one).

3.1 B.

3.2 A.

3.3 (a) users' movements could be tracked or exposed. (b) biased or unfair results. (c) collect only what is needed, secure it, and get consent (any one).