CM50266 Case Study 1 – Answer Guidance

The answers given here are not exhaustive but are a guide to the types of answers expected. Other valid answers may be given and should receive appropriate credit. The goal is to reward understanding of the topic. Good answers will be specific to the question asked and reference aspects of it. Answers that appear generic/general in nature are less likely to be valid answers to the specific question asked.

Q1

The description of each principle does not need to be long, but it should indicate that your peer understands the underlying idea. The wording should be their own and not be cut and pasted from their reference material.

1. Lawfulness, fairness and transparency

Data collected can only be used in lawful ways and the way in which the data is used must be clearly indicated to the user. It should not be used in a way that is unexpected, misleads the user or is unreasonably negative for the user.

2. Purpose limitation

It should only be used for explicit, legal purposes, either where consent has been sought or there is an obligation to do so.

3. Data minimisation

Only the minimum data necessary for the stated purpose should be collected and processed.

4. Accuracy

The data should be up to date and correct. The user should have the means to provide updates and the data controller should actively work to keep data up to date.

5. Storage limitation

Data should only be kept for as long as it is required for the purpose for which it was collected. A specific period need not be specified, but there should be a clear retention policy.

6. Integrity and confidentiality (security)

Data controllers and processors must put in place plans and policies to ensure the data is not compromised either accidentally or deliberately.

Q2

The most obvious example is that the US website relies on an opt-out from marketing. Under UK rules explicit permission in required. The issue can be resolved by changing the opt-out to opt in, provided it is made clear what is being requested.

Q3

There are lots of options to choose from, including,

Maintaining evidence of the steps taken to comply with the UK GDPR.

Maintain ongoing records.

Appoint someone appropriately senior to be responsible for this.

Review the measures in place at regular intervals.

Applying a ‘data protection by design and default’ approach.

Q4

We are not told the nature of the rating that users provide, only that there is also a comment. We can reasonably assume that the rating is either a numeric value on some scale or can be turned into one. There are two related issues here. The use of the rating of a user to provide new information to them and the use of that data to provide new information to others. In both cases this involves the company using the data they already have for a new purpose. They cannot do this unless they obtain explicit permission from the user. However, we would not expect them to be disclosing the data of one user to another directly.

Q5

This can be argued either way depending on what data is preserved and the form in which it is kept.

If the company is adopting best practice, it will have deployed pseudonymisation to separate the clearly PII information such as name, email, etc, from the review data. In theory, removing that PII information could make the remaining data anonymous, however the company must be careful that it is not still just pseudonymised. For example, information in comments left by a user may still allow them to be identified either directly or when combined with other available information sources. The company can potentially retain just the individual rating scores of a user, but the company would need to be confident that they had made them genuinely anonymous. However, where the data has been aggregated into a single value with other users’ data, for example to produce an overall score for a movie, there is no need to unpick it.

You are not awarding marks for a yes/no answer, but the reasoning why that is the answer. Award all three marks for a properly justified answer in Q5a. You may award one mark, if an attempt at justifying has been made, but is weak. For example, if it is claimed that it is sufficient because the ratings are not personally identifying information and therefore not covered but where the fact a user may have put PII within the free form comment is not considered.

Q6

As avatars will be created for existing users, if those were based on the personally identifying data of the users, then it would be making use of the data in a way for which permission has not been granted. But we are asked to put that issue aside and identify another. The problem with using the data as the basis for the generation of the avatar is that the data is then shared with other users, breaching the confidentiality principle. Even if an attempt is made to encrypt or hash it, rainbow tables could potentially be used to ‘reverse-engineer’ the original data.

Q7

The company can use any data that is not personally identifying. This could be either a set of features each chosen at random, or a seed that is used to control the generation. The expectation of uniqueness might make us tempted to use some derivative of a userid, but this could be problematic, if the userid is retained after other PII is removed later. For example, as part of retaining rating information.

Again, it may be tempting to use some form of hashed value based on the user data. This still comes with the risk of exposing user data as in theory inputs could be tried to produce outputs and an understanding of the hash if it’s too simple or a ‘rainbow table’ produced.

Formatting

To receive the mark for formatting, the text should be structured into paragraphs with sections that assist the reader in identifying which part of the document relate to which answers. It is not necessary to strictly follow the questions numbering. For example, it is reasonable to use a series of heading such as ‘Data Protection Principles’, ‘Recommendations’ and ‘Avatars.’

The Bath Harvard referencing style is primarily concerned with the formatting of each entry. To get the mark, most of the references provided must follow the style. References should either be gathered in a list at the end or provided as footnotes. They should not be written in full, with in the text where they are used.