# CM50266 Applied Data Science 2021-2022

Case Study 1 – Data Protection

December 30, 2021

**Abstract**

ACME Review Inc have built a successful online business providing a car review website. It collects the reviews from users and uses these to create overall scores for cars. This American company has decided to expand to the United Kingdom. To do this, they will need to comply with GDPR UK.

Its main source of income is paid-for advertising that is displayed alongside the car review details. When a user signs up, it asks for explicit permission to use information provided by the user to generate ratings for cars and to share individual review comments with site visitors. In addition to allowing the website user to pick a 'screen name' it also requires their real name, age, sex, city of residence and email address. This information is currently used along with the car rating information to help target advertising to individual users. It never shares this information with the advertisers and provides the user the ability to opt-out of the use of the data to select the adverts shown.

## 1 Excluding accountability, what are the data privacy principles of the GDPR?

The UK GDPR (ICO, Accessed Dec 2021a) sets out seven key principles from articles 5(1) and 5(2). It states that personal data should follow:

1. *Lawfulness, fairness and transparency* - The data should be processed within the bounds of the law, in a fair manner and with a level of transparency relating to individuals.

2. *Purpose limitation* - The data should only be collected for legitimate purposes and not processed in such a way which breaks that initial purpose. Processing for archiving purposes which pertain to the public interest, scientific or historical research is not considered to be a breach of the initial purpose.

3. *Data minimisation* - The data should be precisely relevant and limited to match the intentions of the purpose limitation.

4. *Accuracy* - The data should be accurate and kept up to date. Steps must be taken to ensure that inaccuracies are erased or rectified before they are processed.

5. *Storage limitation* - The data should be kept in a form which permits the identification of relevant data for no longer than is necessary for the processing steps. Personal data may be stored for longer periods only if the intention is for archive purposes or scientific research, subject to the measures required to safeguard to rights and freedoms of individuals.

6. *Integrity and confidentiality* (security) - The data should be processed in such a way that ensures its security of personal data, including that which is deemed unauthorised and unlawful processing against accidental loss, destruction or damage.

7. *Accountability* - The controller is entirely responsible to demonstrate compliance with all of the previous points.

## 2 Identify a change to the way the current US website works that the company will need to make to be compatible with the GDPR when it launches the UK version, and why this is necessary.

As we can assume that Acme have decided to keep the UK and US sites entirely separate, with the data collected in the UK they can avoid content sharing issues which cannot be shown in the US (BBC, Accessed Dec 2021). UK personal user data must be stored in accordance with GDPR (GOV, Accessed Dec 2021) and remain securely in its country of origin.

The one change they need to make in their codebase is to implement a customer consent form from visitors in the UK and EU (CookieBot, Accessed Dec 2021). The explicit consent of users must be obtained before any processing or transfer is allowed to take place. Even though the UK left the EU under Brexit, we have a new domestic data privacy law called UK-GDPR which is exactly the same as the EU version, this compliance within the UK is still an obligation for websites in this domain.

# 3 Indicate two actions the company will need to take in relation to the implementation of the new features described above, because of the GDPR Accountability principle.

The company plans to introduce two new features:

1. It will recommend cars it thinks a user will like based on their having a similar profile of existing car ratings to other site users.

2. It currently does not have any kind of avatar image that is displayed when comments are shown. It will allow users to upload a small image to be used for this purpose. If a user has not uploaded an image, it plans to assign a system generated cartoon 'face'. The construction of the face will be based on the additional information that users have previously provided to the company. The company believes this will be more effective than using a single standard blank face cartoon. It will also allow users to modify this cartoon. For example, by changing hair style, eye colour, add glasses and similar features.

Starting with the first implementation, the company will need to most importantly implement[1] from the GDPR (ICO, Accessed Dec 2021a):

1. *Purpose Limitation/Data Minimisation* - To recommend cars it thinks a user will like, matching user data needs to be processed within the bounds of the initial purpose and design of the recommender.

2. *Integrity and confidentiality* - As personal data is required to match ratings for the recommender, this should be processed in such a way that ensures the security of the data therein.

The second implementation will require the company to most importantly implement:

1. *Integrity and confidentiality* - As personal data is required to generate a cartoon face if a user does not upload an image, this should be processed in such a way that ensures the security of the data therein. Consent to do so would also need to be provided, and issues such as gender/race bias need to be carefully investigated and considered.

2. *Accuracy* - The personal data to generate cartoon profile pictures needs verification steps to ensure accuracy is maintained in order to be most effective.

[1] The company needs to implement all the GDPR Accountability principles (GOV, Accessed Dec 2021).

## 4 Identify a GDPR related issue that the company may have with implementing the plan to provide individualised recommendations and suggest a way these could be addressed to allow this to proceed.

The main issue relating to UK-GDPR in the collecting of personal information which requires a cookie consent form to be authorised by visitors to a UK site CookieBot (Accessed Dec 2021). Once this has been implemented, strict guidance needs to be adhered to according to the GDPR Accountability Principle.

Of immediate importance relating to data collection is *Storage Limitation* (ICO, Accessed Dec 2021a). The data should be kept in a form which permits the identification of relevant data for no longer than is necessary for the processing steps. In our case, we need only temporarily store user correlated data to match reviews for the recommender, once it has been delivered, there is no need to store the information permanently.

## 5 When a user decides to close their account on the website, the company is required to delete their data. In order to continue to provide the useful ratings and review comments to other users, the company would like to turn this data into anonymous data by disconnecting it from the personal details (name, city, etc.) held about the user. It plans to seek permission to do this. Is the deleting of the personal data sufficient to achieve this? Explain why it is/is not sufficient.

The UK GDPR introduces a right for individuals to have personal data erased (ICO, Accessed Dec 2021b) and is defined by:

1. The right to erasure is also known as 'the right to be forgotten'.
2. The right is not absolute and only applies in certain circumstances.
3. Individuals can make a request for erasure verbally or in writing.
4. You have one month to respond to a request.
5. This right is not the only way in which the UK GDPR places an obligation on you to consider whether to delete personal data.

Individuals have the right to have their personal data erased (Consulting, Accessed Dec 2021) if:

1. the personal data is no longer necessary for the purpose which you originally collected or processed it for;
2. you are relying on consent as your lawful basis for holding the data, and the individual withdraws their consent;
3. you are relying on legitimate interests as your basis for processing, the individual objects to the processing of their data, and there is no overriding legitimate interest to continue this processing;
4. you are processing the personal data for direct marketing purposes and the individual objects to that processing;
5. you have processed the personal data unlawfully (ie in breach of the lawfulness requirement of the 1st principle);

6. you have to do it to comply with a legal obligation; or

7. you have processed the personal data to offer information society services to a child.

ACME wishes to make users anonymous upon deletion of their user accounts by retaining review details, whilst erasing personal details such as name, address etc.

If the customer requests a full data deletion after consent was given to remove personal details but keep the remaining data, the 'right to be forgotten' can no longer be processed, as the user data would have been anonymised. This is in violation of the UK-GDPR (ICO, Accessed Dec 2021a), as simply relying on prior consent as a lawful basis for holding the data is insufficient, as the individual can withdraw their consent.

# 6 Other than a lack of consent, suggest a reason that allowing the system to generate the avatar image in the way described would not be compatible with the GDPR.

Suppose we have user details such as age, gender and location and we generate a cartoon avatar with consent from a user. The cartoon image leaks their data by allowing other visitors to infer their details from the image generated. This data leakage would be in violation of GDPR (ICO, Accessed Dec 2021a), and thus generating an avatar image in this way would not be compatible with GDPR.

# 7 Indicate an alternative approach that could be employed to provide a unique system generated avatar image for each user that would be compatible with the GDPR and would not leak any of the user details. And explain why this would be compatible.

An alternative approach could be taken to simply use the initials of the user as the default image setting (phppot, Accessed Dec 2021). Other users would be unable to infer any personal data about a user from initials alone, and this solution would be compatible with UK-GDPR.

Adding the feature to generate an initial avatar over the cartoon image has the following advantages:

1. It gives a good user experience by generating avatars automatically instead of forcing them to choose a profile photo.
2. Application-specific default avatars may make the user's profile page with the position of anonymity. This will be prevented by generating an initial avatar from names.
3. Initials are the most probably used icons for the person's identity in real-time. So this concept aptly suits for this scenario.
4. The best and simple solution to personalize the user's profile page.
5. Reduces users time and effort by simplifying the process of entering the registration and profile update form.

Alternatively, it would be possible to create a model based on (Medium, Accessed Dec 2021), which would allow a user to upload an image to create a bitmoji cartoon. This would allow for the initial design to be adapted to be in line with the GDPR.

# 8 References


BBC (Accessed Dec 2021) "Why is the BBC website address changing from bbc.co.uk to bbc.com?," https://www.bbc.co.uk/contact/questions/behind-the-scenes/website-domain-change.

Consulting, Intersoft (Accessed Dec 2021) "Using personal data in your business or other organisation," https://gdpr-info.eu/art-17-gdpr/.

CookieBot (Accessed Dec 2021) "Using personal data in your business or other organisation," https://www.cookiebot.com/en/gdpr-in-the-uk-2021-uk-adequacy-decision-update/?gclid=Cj0KCQiAq7COBhC2ARIsANsPATHL0tzwzwDfT2mJB3brg-eTajTQXuer7o7wPENHBTFeYMGd_zyGem8aAmMIEALw_wcB.

GOV (Accessed Dec 2021) "Using personal data in your business or other organisation," https://www.gov.uk/guidance/using-personal-data-in-your-business-or-other-organisation.

ICO (Accessed Dec 2021a) "Guide to the General Data Protection Regulation (GDPR)," https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/.

——— (Accessed Dec 2021b) "Using personal data in your business or other organisation," https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-erasure/.

Medium (Accessed Dec 2021) "Using personal data in your business or other organisation," https://towardsdatascience.com/avatargan-generate-cartoon-images-using-gan-1ffe7d33cfbb.

phppot (Accessed Dec 2021) "Using personal data in your business or other organisation," https://phppot.com/php/how-to-generate-initial-avatar-image-from-username-using-php-dynamically/.