Question 3

29 marks

A state government agency launches an open data platform making various public sector datasets available to citizens, researchers, and businesses. One particular dataset includes records of public service usage (e.g., library loans, park facility bookings) which initially contained some direct personal identifiers. Before release, the agency undertook a process to remove these identifiers and generalise other data points. The platform also includes a feedback mechanism where users can report data inaccuracies.

a) Explain the process and significance of de-identification for data being released on a public open data platform. Discuss the ethical practice of purpose limitation in this context, and how it relates to the initial collection of personal identifiers. (10 marks)

b) Discuss the critical role of data quality for datasets released on an open data platform and explain how the government agency can ensure the information is accurate, up-to-date, and complete. Furthermore, explain the importance of transparency regarding the data processing methods used before public release. Identify the relevant Australian Privacy Principles for these aspects.
(9 marks)

c) Analyse how the feedback mechanism allowing users to report data inaccuracies aligns with an individual's right to have their personal information corrected. Consider a scenario where the agency receives unsolicited personal information through this feedback channel. Identify the relevant Australian Privacy Principles and describe the agency's obligations in both these situations. (10 marks)

Marking Guide

Part A

Marks	Criteria	Evidence of achievement
1	Defines de-identification	States de-identification removes or generalises direct personal identifiers.
1	Explains purpose of de-identification	Protects individual privacy while retaining data utility.
1	Explains risk without de-identification	Notes risk of re-identification and privacy breach.
1	Gives example in context	Example: removing names, generalising age into bands.
1	Defines purpose limitation	States data should only be used for the purpose originally collected.
1	Links purpose limitation to ethics	Notes it prevents function creep or misuse of personal identifiers.
1	Connects to open data release	Explains identifiers removed because not required for open data purpose.
1	Identifies APP 6 (Use/disclosure)	APP 6 limits use/disclosure to original purpose or with consent.
1	Explains APP 6 in context	Example: identifiers collected for service delivery must not be exposed in open data.
1	Integrates APP with ethical practice	Concludes de-identification and purpose limitation uphold privacy and trust.

Sample Response

De-identification is the process of removing or generalising direct identifiers, such as names, addresses, or account numbers, so individuals cannot be reasonably re-identified. In the case of public service usage records, this might involve replacing names with anonymised IDs and grouping ages into ranges. The significance is that it protects individual privacy while still allowing researchers and businesses to analyse service trends. Without de-identification, releasing the dataset could expose personal details and cause harm.

The ethical principle of purpose limitation requires that data be used only for the purpose for which it was originally collected. For example, personal identifiers were initially collected to manage individual library loans or facility bookings, not for public release. By removing identifiers, the agency ensures that the open data is consistent with the original purpose of collection while still being useful.

APP 6 (Use and disclosure of personal information) supports this practice by restricting disclosure of personal data for secondary purposes without consent. Removing identifiers before public release ensures compliance. De-identification and purpose limitation together balance open data goals with ethical and legal duties to respect privacy.

Part B

Marks	Criteria	Evidence of achievement
1	Defines data quality	States quality = accuracy, completeness, and up-to-dateness.
1	Explains why data quality matters	Inaccurate data reduces usefulness and can mislead researchers/businesses.
1	Provides strategy: accuracy	Example: regular validation against source systems.
1	Provides strategy: up-to-dateness	Example: scheduled updates or real-time syncing.
1	Provides strategy: completeness	Example: ensuring all records are uploaded, not partial.
1	Explains transparency	States agency must disclose how data was processed/de-identified before release.
1	Identifies APP 10 (Quality)	APP 10 requires organisations to keep personal information accurate and complete.
1	Identifies APP 1 (Transparency)	APP 1 requires openness about management of personal data.
1	Explains relevance of APPs	Links APP 10 and APP 1 to maintaining dataset reliability and public trust.

Sample Response

Data quality is essential for open datasets because inaccurate, outdated, or incomplete information reduces their reliability and can mislead researchers or businesses. For example, if park booking records are missing or out of date, users may base decisions on false patterns.

The agency can promote accuracy by validating datasets against source systems before release, ensure up-to-dateness through scheduled refresh cycles, and maintain completeness by confirming that all records, not just subsets, are included.

Transparency is also vital. The agency should explain the methods used for de-identification, aggregation, and generalisation so users can understand limitations. Publishing metadata that describes the processing steps builds trust and helps researchers interpret results correctly.

APP 10 (Quality of personal information) requires personal data to be accurate, complete, and current before use or disclosure, while APP 1 (Open and transparent management) requires organisations to communicate their information-handling processes. Together, these principles ensure datasets are both reliable and responsibly released.

Part C

Marks	Criteria	Evidence of achievement
1	Identifies right to correction	Notes individuals can request correction of inaccurate data.
1	Identifies APP 13 (Correction)	APP 13 gives individuals the right to correction of their personal information.
1	Explains APP 13 in context	Example: feedback mechanism enables reporting of incorrect service usage records.
1	Identifies transparency obligation	Agency must confirm corrections are made and communicate actions to the user.
1	Analyses unsolicited information	Defines unsolicited information as data not actively requested by agency.
1	Identifies APP 4 (Unsolicited info)	APP 4 sets obligations when unsolicited personal information is received.
1	Explains APP 4 in context	Example: if user accidentally submits private identifiers in feedback.
1	States required response	Agency must assess if information could have been lawfully collected; if not, destroy.
1	Integrates correction and unsolicited	Shows both processes protect privacy and uphold accountability.
1	Ethical conclusion	Concludes mechanisms empower users and limit unnecessary data retention.

Sample Response

The feedback mechanism that allows users to report inaccuracies aligns with an individual's right to correction. If an individual believes their public service usage record is wrong, the agency must investigate and update the dataset if necessary.

APP 13 (Correction of personal information) requires agencies to correct inaccurate personal information and to take reasonable steps to ensure corrections are communicated. For example, if a library loan was incorrectly attributed, the agency should amend the record and republish corrected data.

Sometimes, users may submit unsolicited personal information through the feedback channel, such as including their name and phone number in a complaint. APP 4 (Unsolicited personal information) requires the agency to determine whether it could have lawfully collected this data. If not, it must be destroyed or de-identified as soon as practicable.

By upholding APP 13 and APP 4, the agency ensures that the feedback process supports data accuracy while preventing the unnecessary retention of private information. This empowers users to engage with the platform while maintaining strong privacy protections.