PERSONAL DATA PROTECTION COMMISSION SINGAPORE

PRACTICAL GUIDANCE TO SPH MEDIA

Background

1. SPH Media is implementing a Proof of Concept (POC) to enable its partners to find prospective customers from SPH's customer pool in order to serve them relevant advertisements.

2. Key stakeholders involved in the POC are as follows:

   a) SPH Media ("**Publisher**") – Owns the advertising platform where advertisements would be shown

   b) Global Wealth Manager ("**Advertiser**") – Intends to increase advertising-effectiveness by tapping on SPH Media's customer database to find new customers.

   c) Data Management Platform ("**DMP**") – Partners SPH Media to ensure that the relevant advertisement is shown to the right audience.

   d) Decentriq ("**Solution Provider**") – Provider of PETs solution built on Trusted Execution Enclave ("**TEE**") technology

3. Key steps involved in the POC include:

   a) **Creating list of lookalike customers within the Trusted Execution Enclave (TEE).**

      i. **Hashing of identifiers.** Publisher and Advertiser will each upload their customer information directly into the TEE through an encrypted channel. The agreed customer identifier (i.e., email addresses) in the customer information will be hashed using the same hash algorithm. The Advertiser's customer information includes a list of its customers' hashed email addresses, while the Publisher's customer information includes a list of its customers' hashed email addresses, corresponding generated SPH User ID ("SPH ID") and corresponding segment IDs which represent the interest areas of each customer (e.g., segment ID 001 to represent soccer, ID 002 for fashion etc).

      ii. **Creation of seed list and lookalike audience.** The two lists of hashed emails uploaded by Publisher and Advertiser will be matched against each other within the TEE to generate the list of common

customers which will form a “seed list”. Further analysis of this seed list against the rest of Publisher’s customer information will be done within the TEE to generate a list of “lookalike customers which have similar profiles as customers in the “seed list”. The list of “lookalike customers” will comprise a list of the customers’ SPH ID and the corresponding segment ID. The Advertiser will not have access to the seed list nor the “lookalike customers” list.

b) **Serving advertisements to lookalike customers.**

i. **Sharing of lookalike customer list with DMP.** The Publisher will share the list of “lookalike customers” comprising the list of SPH IDs and the corresponding customer segment IDs with the DMP. The SPH IDs will expire after 30 days from its creation.

ii. **Transformation of customer segment ID in lookalike customer list.** The DMP will transform the customer segment ID to a DMP-segment ID. The Publisher-segment ID and DMP-segment ID mapping will be stored by the DMP.

iii. **Serving of advertisements when lookalike customer lands on Publisher’s website.** Each time a user lands on the Publisher’s website, the DMP (through the embedded script on the website) will retrieve the user’s SPH IDs from the cookie on the user’s browser and cross-check against the SPH IDs in the list of lookalike customers (see para 3(b)(i) above). If the user possesses the SPH ID that matches the lookalike list, the DMP will drop the corresponding DMP-segment ID into the local storage of the browser. The DMP-segment ID will be picked up by **SPH Media’s Publisher-side Ad Server (“PAS”)** script on the website which will then serve the relevant advertisements targeting the specific segment ID.

iv. **Deletion of SPH IDs and Segment IDs.** Upon conclusion of the advertising campaign, the DMP will delete the “lookalike customer” list (i.e., SPH IDs and Segment IDs) that had been shared by the Publisher.

4. SPH Media sought Practical Guidance (Guidance) from the Personal Data Protection Commission (PDPC) on the following:

a) Whether the Publisher and Advertiser would be considered to have disclosed personal data to each other by uploading of their respective list of customer information for processing within the TEE; and

b) Whether the SPH IDs and customer segment IDs that the Publisher shares with the DMP partner constitutes personal data under the PDPA.

PDPC's assessment

Whether the Publisher and Advertiser have disclosed personal data for processing within the TEE

5. In this POC, PDPC notes that the solution is designed such that the Publisher and Advertiser will not be able to access each other's data input to the TEE. Given so, PDPC is of the view that there is likely to be no disclosure of personal data between the Publisher and Advertiser. In particular, we note that the following safeguards have been implemented in the POC to prevent access to the data uploaded to the TEE:

a) Hashing. Email addresses uploaded into the TEE for matching will be hashed with SHA-256. SHA-256 is a specified secure hashing algorithm under NIST FIPS 180-4[1] which adds an additional layer of protection from unauthorised disclosure/access by third parties

b) Implementation of TEE solution. The use of TEE solution ensures that data is executed in a secure computing environment and inaccessible to all parties (i.e., Advertiser, Publisher, TEE solution provider), and that only authorised codes/algorithms can be run (and subsequently verified) within the TEE. It leverages technology for hardware-based isolation and attestation to ensure data and code integrity even if the main operating system is compromised.

6. Nevertheless, the Publisher and Advertiser would be considered to have used personal data by uploading their hashed email lists to the TEE and generating both the seed list and the lookalike customer list, and the Data Protection Provisions under the PDPA will apply (e.g., Consent Obligation, Protection Obligation). As part of this POC, both the Advertiser and Publisher may consider relying on PDPA's Business Improvement Exception (BIE) to use its customers' personal data without consent given that the intent is (i) to learn and understand its customers preferences and (ii) to personalise relevant goods and services for the users in the "lookalike" list.

Whether the SPH IDs and customer segment IDs that the Publisher shares with the DMP constitutes personal data under the PDPA.

7. PDPC is of the view that the SPH IDs and customer segment IDs shared by the Publisher with the DMP will not likely constitute personal data under the PDPA due to the very low risk of re-identification of individuals. While unique to individuals, both the SPH ID and Segment ID are indirect identifiers created by the Publisher which will be meaningful only to the Publisher. The DMP is unlikely to have the ability to link these identifiers to its own customer records without additional information about this individual. In addition, PDPC notes that SPH IDs are temporary IDs related to individuals, which will expire after 30 days from its creation.

8. To reduce the likelihood of re-identification, the DMP partner has implemented an additional transformation step to convert the Publisher-segment ID into DMP-

[1] NIST FIPS 180-4 Secure Hash Standard specifies hash algorithms that can be used to generate digests of messages. The digests are used to detect whether messages have been changed since the digests were generated.

segment ID before sharing with the PAS. This prevents the exposure of the Publisher-segment ID to additional parties, which in turn reduces the likelihood of the Publisher-segment ID from being linked or associated with specific individuals where more information may be gleaned from such categorisation of profiles/segments.

Additional safeguards that can be implemented

9. PDPC is of the view that there are additional safeguards that Parties may wish to consider implementing as part of this POC:

a) Prior to data processing within the TEE, Parties may consider using a shared common salt in the hashing process to prevent rainbow table attacks (pre-computed hash tables). Parties should also ensure they are using current implementations of SHA-256 from well-maintained cryptographic libraries and regularly update these libraries to incorporate security patches and implementation improvements.

b) While TEE solutions offer secure infrastructure, organisations that engage third party TEE solution providers remain responsible under the PDPA for the data processing within the TEE. Both the Publisher and Advertiser may wish to conduct due diligence to assess and validate if the data security afforded by TEE solution, and the data governance policies and practices of the solution provider are sufficient and appropriate for their business needs.

END OF DOCUMENT