metadata of the received data; identifying, by the inflow data analyzer, a search space comprising at least one sensitive attribute from the processed metadata based on a pre-defined knowledge base associated with the data, wherein the at least one sensitive attribute comprises one or more of binary, categorical, numerical and descriptive texts; generating, by an adversary model generator, an adversary model by partitioning the search space into sets of buckets, each set corresponding to the at least one sensitive attribute having a privacy data associated thereof; computing, by the adversary model generator, bucket count for each of the sets of buckets and creating bucket combinations of the buckets from the sets of buckets; replacing, by the adversary model generator, the privacy data associated with each of the at least one sensitive attribute with a masking bucket from the buckets; computing, by the adversary model generator, an entity count for each of the bucket combinations based on the masking bucket assigned to the privacy data; computing, by the adversary model generator, an anonymity index based on the computed entity count and a pre-defined privacy threshold; and sanitizing, by a data masking module, the privacy data based on the computed anonymity index to generate output data.

[009] In another aspect, there is provided a system comprising: one or more processors; and one or more internal data storage devices operatively coupled to the one or more processors for storing instructions configured for execution by the one or more processors, the instructions being comprised in: a data connector configured to: receive data from one or more data sources for making the data consumable by one or more data buyers; an inflow data analyzer configured to: analyze the received data to extract and process metadata of the received data; and identify a search space comprising at least one sensitive attribute from the processed metadata based on a pre-defined knowledge base associated with the data; an adversary model generator configured to: generate an adversary model by partitioning the search space into sets of buckets, each set corresponding to the at least one sensitive attribute having a privacy data associated thereof; wherein the at least one sensitive attribute comprises one or more of binary, categorical, numerical and descriptive texts, compute bucket count for each of the sets of buckets and creating bucket combinations of the buckets from the sets of buckets; replace the privacy data associated with each of the at least one sensitive attribute with a masking bucket from the buckets; compute an entity count for each of the bucket combinations based on the masking bucket assigned to the privacy data; compute an anonymity index based on the computed entity count and a pre-defined privacy threshold; and continually learn and update the adversary model based on the received data; a data masking module configured to: sanitize the privacy data based on the computed anonymity index to generate output data; a decision helper module configured to: provide recommendations to data sellers based on the received data; an outflow

WHAT IS CLAIMED IS:

1. A processor implemented method (200) comprising:

receiving, by a data connector, data from one or more data sources for making the data consumable by one or more data buyers (202);

analyzing, by an inflow data analyzer, the received data to extract and process metadata of the received data (204);

identifying, by the inflow data analyzer, a search space comprising at least one sensitive attribute from the processed metadata based on a pre-defined knowledge base associated with the data, wherein the at least one sensitive attribute comprises one or more of binary, categorical, numerical and descriptive texts (206);

generating, by an adversary model generator, an adversary model by partitioning the search space into sets of buckets, each set corresponding to the at least one sensitive attribute having a privacy data associated thereof (208);

computing, by the adversary model generator, bucket count for each of the sets of buckets and creating bucket combinations of the buckets from the sets of buckets (210);

replacing, by the adversary model generator, the privacy data associated with each of the at least one sensitive attribute with a masking bucket from the buckets (212);

computing, by the adversary model generator, an entity count for each of the bucket combinations based on the masking bucket assigned to the privacy data (214);

computing, by the adversary model generator, an anonymity index based on the computed entity count and a pre-defined privacy threshold (216); and

sanitizing, by a data masking module, the privacy data based on the computed anonymity index to generate output data (218).

2. The processor implemented method of claim 1, wherein generating an adversary model comprises generating the sets of buckets based on one of (a) pre-defined ranges of values or (b) pre-defined upper and lower bounds, for each of the at least one sensitive attribute based on the pre-defined knowledge base.

3. The processor implemented method of claim 2, wherein the range of values and the upper and lower bounds are computed by the inflow data analyzer based on the at least one sensitive attribute.

4. The processor implemented method of claim 1, wherein sanitizing the privacy data comprises one of (i) hierarchy masking techniques, (ii) bucket masking techniques, (iii) clustering technique or (iv) shuffling technique, wherein the shuffling technique is performed to obtain balanced buckets, wherein each bucket combination has a balanced entity count.

5. The processor implemented method of claim 5 further comprising computing, by a data privacy-utility tradeoff calculator, a utility index based on mid-point of the balanced bucket and the privacy data (220).

6. The processor implemented method of claim 6 further comprising computing, by the data privacy-utility tradeoff calculator, attribute variations based on the number of variations between the buckets and the balanced buckets (222).

7. The processor implemented method of claim 1 further comprises continually learning and updating the adversary model based on the received data (224).

8. The processor implemented method of claim 7 further comprising one or more of:

providing recommendations, by a decision helper module, to data sellers based on the received data (226);

evaluating, by an outflow data analyzer, the output data to match requirements of the one or more data buyers (228); and

determining, by a data release management module, a release plan based on the recommendations by the decision helper module (230);

generating, by a report and alert management module, evaluation reports and alerts based on the output data (232); and

logging, by an event logging module, events associated with the output data (234).

9. A system (100) comprising:

one or more processors (102); and

one or more internal data storage devices (106) operatively coupled to the one or more processors (102) for storing instructions configured for execution by the one or more processors (102), the instructions being comprised in:

a data connector (108A) configured to:

receive data from one or more data sources for making the data consumable by one or more data buyers;

an inflow data analyzer (108B) configured to:

analyze the received data to extract and process metadata of the received data; and

identify a search space comprising at least one sensitive attribute from the processed metadata based on a pre-defined knowledge base associated with the data;

an adversary model generator (108C) configured to:

generate an adversary model by partitioning the search space into sets of buckets, each set corresponding to the at least one sensitive attribute having a privacy data associated thereof, wherein the at least one sensitive attribute comprises one or more of binary, categorical, numerical and descriptive texts;

compute bucket count for each of the sets of buckets and creating bucket combinations of the buckets from the sets of buckets;

replace the privacy data associated with each of the at least one sensitive attribute with a masking bucket from the buckets;

compute an entity count for each of the bucket combinations based on the masking bucket assigned to the privacy data;

compute an anonymity index based on the computed entity count and a pre-defined privacy threshold; and

continually learn and update the adversary model based on the received data;

a data masking module (108D) configured to:

sanitize the privacy data based on the computed anonymity index to generate output data;

a decision helper module (108F) configured to:

provide recommendations to data sellers based on the received data;

an outflow data analyzer (108G) configured to:

evaluate the output data to match requirements of the one or more data buyers;

a data release management module (108H) configured to:

determine a release plan based on the recommendations by the decision helper module (108F);

a report and alert management module (108I) configured to:

generate evaluation reports and alerts based on the output data; and

an event logging module (108J) configured to:

log events associated with the output data.

10. The system of claim, wherein the adversary model generator 11 (108C) is further configured to generate the sets of buckets based on one of (a) pre-defined ranges of values or (b) pre-defined upper and lower bounds, for each of the at least one sensitive attribute based on the pre-defined knowledge base.

11. The system of claim 10, wherein the range of values and the upper and lower bounds are computed by the inflow data analyzer (108B) based on the at least one sensitive attribute.

12. The system of claim 9, wherein the data masking module (108D) is further configured to sanitize the privacy data by one of (i) hierarchy masking techniques, (ii) bucket masking techniques, (iii) clustering technique or (iv) shuffling technique, wherein the shuffling technique is performed to obtain balanced buckets, wherein each bucket combination has a balanced entity count.

13. The system of claim 11 further comprising:

a data privacy-utility tradeoff calculator (108E) configured to:

compute a utility index based on mid-point of the balanced buckets and the privacy data; and

compute attribute variations based on the number of variations between the buckets and the balanced buckets.

14. A computer program product comprising a non-transitory computer readable medium having a computer readable program embodied therein, wherein the computer readable program, when executed on a computing device, causes the computing device to:

receive data from one or more data sources for making the data consumable by one or more data buyers;

analyze the received data to extract and process metadata of the received data;

identify a search space comprising at least one sensitive attribute from the processed metadata based on a pre-defined knowledge base associated with the data;

generate an adversary model by partitioning the search space into sets of buckets, each set corresponding to the at least one sensitive attribute having a privacy data associated thereof;

compute a bucket count for each of the sets of buckets and creating bucket combinations of the buckets from the sets of buckets;

replace the privacy data associated with each of the at least one sensitive attribute with a masking bucket from the buckets;

compute an entity count for each of the bucket combinations based on the masking bucket assigned to the privacy data;

compute an anonymity index based on the computed entity count and a pre-defined privacy threshold; and

sanitize the privacy data based on the computed anonymity index to generate output data and obtain balanced buckets, wherein each bucket combination has a balanced entity count.

15. The computer program product of claim 16, wherein the computer readable program further causes the computing device to perform one or more of:

compute a utility index based on mid-point of the balanced bucket and the privacy data;

compute attribute variations based on the number of variations between the buckets and the balanced buckets;

continually learn and update the adversary model based on the received data;

provide recommendations to data sellers based on the received data;

evaluate the output data to match requirements of the one or more data buyers;

determine a release plan based on the provided recommendations;

generate evaluation reports and alerts based on the output data; and

log events associated with the output data.