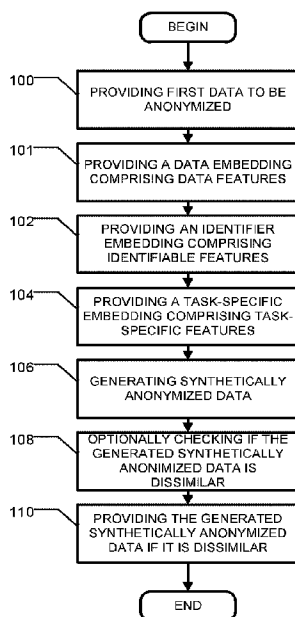




(86) Date de dépôt PCT/PCT Filing Date: 2019/07/12  
(87) Date publication PCT/PCT Publication Date: 2020/01/16  
(45) Date de délivrance/Issue Date: 2023/08/22  
(85) Entrée phase nationale/National Entry: 2020/12/31  
(86) N° demande PCT/PCT Application No.: IB 2019/055972  
(87) N° publication PCT/PCT Publication No.: 2020/012439  
(30) Priorité/Priority: 2018/07/13 (US62/697,804)

(51) Cl.Int./Int.Cl. *G06F 21/60* (2013.01),  
*G16H 10/60* (2018.01)  
(72) Inventeurs/Inventors:  
CHANDELIER, FLORENT, CA;  
JESSON, ANDREW, CA;  
DIJORIO, LISA, CA;  
LOW-KAM, CECILE, CA;  
SOUDAN, FLORIAN, CA;  
HAVAEI, MOHAMMAD, CA;  
CHAPADOS, NICOLAS, CA  
(73) Propriétaire/Owner:  
IMAGIA CYBERNETICS INC., CA  
(74) Agent: FASKEN MARTINEAU DUMOULIN LLP

(54) Titre : PROCEDE ET SYSTEME DE GENERATION DE DONNEES SYNTHETIQUEMENT ANONYMISEES POUR UNE TACHE DONNEE  
(54) Title: METHOD AND SYSTEM FOR GENERATING SYNTHETICALLY ANONYMIZED DATA FOR A GIVEN TASK



(57) Abrégé/Abstract:

A method and a system are disclosed for generating synthetically anonymized data, the method comprising providing first data to be anonymized; providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data; providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data; providing a task-specific embedding comprising task-specific features, wherein said task-specific features enables a disentanglement of different classes relevant to the given task; generating synthetically anonymized data, the generating comprising a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process.

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

16 January 2020 (16.01.2020)



(10) International Publication Number

WO 2020/012439 A1

(51) International Patent Classification:

G06F 21/60 (2013.01) G16H 10/60 (2018.01)

(21) International Application Number:

PCT/IB2019/055972

(22) International Filing Date:

12 July 2019 (12.07.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/697,804 13 July 2018 (13.07.2018) US

(71) Applicant: **IMAGIA CYBERNETICS INC.** [CA/CA];  
6650, rue St-Urbain, suite 100, Montréal, Québec H2S 3G9  
(CA).

(72) Inventors: **CHANDELIER, Florent**; 1245 chemin du Lac  
Saint-Louis, Lery, Québec J6N 1A9 (CA). **JESSON, An-**  
**drew**; 802, Avenue Dollard, Apt. 5, Outremont, Québec  
H2V 337 (CA). **HAVAEI, Mohammad**; 4-2945 chemin  
Bedford, Montréal, Québec H3S 1G3 (CA). **DIJORIO,**  
**Lisa**; 4593 Cartier, Montréal, Québec H2H 1W9 (CA).  
**LOW-KAM, Cecile**; 4593 Cartier, Montréal, Québec  
H2H 1W9 (CA). **CHAPADOS, Nicolas**; 7081 Christophe  
Colomb, Montréal, Québec H2S 2H4 (CA). **SOUDAN, Flo-**

rian; 5838, av de Chateaubriand, Montréal, Québec H2S  
2N2 (CA).

(74) Agent: **FASKEN MARTINEAU DUMOULIN**  
**S.E.N.C.R.L., S.R.L.**; 800, rue du Square-Victoria, bureau  
3700, Montréal, Québec H4Z 1E9 (CA).

(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,  
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,  
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,  
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,  
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,  
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,  
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,  
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,  
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every  
kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,  
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,  
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,  
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,  
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: METHOD AND SYSTEM FOR GENERATING SYNTHETICALLY ANONYMIZED DATA FOR A GIVEN TASK

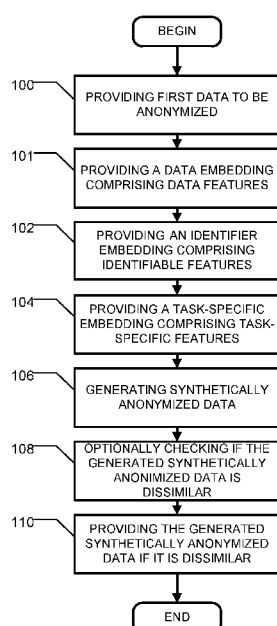


FIG. 1

(57) Abstract: A method and a system are disclosed for generating synthetically anonymized data, the method comprising providing first data to be anonymized; providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data; providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data; providing a task-specific embedding comprising task-specific features, wherein said task-specific features enables a disentanglement of different classes relevant to the given task; generating synthetically anonymized data, the generating comprising a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process.

# WO 2020/012439 A1

---

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**Published:**

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

## METHOD AND SYSTEM FOR GENERATING SYNTHETICALLY ANONYMIZED DATA FOR A GIVEN TASK

### TECHNICAL FIELD

5 The invention relates to data processing. More precisely, the invention pertains to a method and system for generating synthetically anonymized data for a given task.

### BACKGROUND

Being able to provide anonymized data is of great interest for various reasons.

10 Recently, AI methods have been introduced as part of the Statistical methods protecting sensitive information or the identity of the data owner have become critical to ensure privacy of individuals as well as of organizations.

Specifically, sharing individual-level data from clinical studies remains challenging. The status quo often requires scientists to establish a formal collaboration and execute extensive data usage agreements before sharing data. These requirements slow or even prevent data sharing between researchers in all but the closest  
15 collaborations and are serious drawbacks.

Recent initiatives have begun to address cultural challenges around data sharing. In recent years, many datasets containing sensitive information about individuals have been released into public domain with the goal of facilitating data mining research. Databases are frequently anonymized by simply suppressing identifiers that reveal  
20 the identities of the users, like names or identity numbers.

Different processes (<https://arxiv.org/pdf/1802.09386.pdf>; <https://arxiv.org/pdf/1803.11556.pdf>; <https://www.biorxiv.org/content/biorxiv/early/2017/07/05/159756.full.pdf>; <https://openreview.net/forum?id=rJv4XWZA->) are of great value in the anonymization  
25 process of data to either augment training data (See Synthetic data augmentation using GAN for improved liver lesion classification

[http://www.eng.biu.ac.il/goldebej/files/2018/01/ISBI\\_2018\\_Maayan.pdf](http://www.eng.biu.ac.il/goldebej/files/2018/01/ISBI_2018_Maayan.pdf)) or share subject data, however they do not feature the following two requirements: (1) a guarantee that the generated data is not identifiable (background attacks, including attacks if you know, a posteriori, tasks for which the anonymized data was well suited for), and (2) a guarantee that the generated data is relevant for a subsequent task (disentangling appropriate factors of task-specific variations).

There is a need for a method and system that will overcome at least one of the above-identified drawbacks.

Features of the invention will be apparent from review of the disclosure, drawings and description of the invention below.

### **BRIEF SUMMARY**

According to a broad aspect, there is disclosed a method for generating synthetically anonymized data for a given task, the method comprising providing first data to be anonymized; providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data; providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data; providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enable a disentanglement of different classes relevant to the given task; generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the

generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data; and providing the generated synthetically anonymized data for the given task.

5 In accordance with an embodiment, the generating of the synthetically anonymized data for the given task comprises checking that the synthetically anonymized data is dissimilar to the first data to be anonymized for a given metric and the generated synthetically anonymized data for the given task is provided if said checking is successful.

According to an embodiment, the first data comprises patient data.

10 According to an embodiment, the providing of the task-specific embedding comprising task specific features suitable for said task comprises obtaining an indication of the given task; obtaining an indication of classes relevant to the given task; obtaining a model suitable for performing a disentanglement of the data for the given task; and  
15 generating the task-specific embedding using the obtained model, the indication of classes relevant to the given task, the indication of the given task and the data.

According to an embodiment, the providing of the identifier embedding comprising identifiable features comprises obtaining data used for identifying the identifiable features; obtaining a model suitable for identifying the identifiable features in said data; obtaining an indication of identifiable entities and generating the identifier  
20 embedding using the model suitable for identifying the identifiable features, the indication of identifiable entities and the data to be used for identifying the identifiable features.

According to an embodiment, the data comprises the data used for identifying the identifiable features.

25 According to an embodiment, the model suitable for identifying the identifiable features in the data comprises a Single Shot MultiBox Detector (SSD) model.

According to an embodiment, the model suitable for performing a disentanglement of the data for the given task comprises one of an Adversarially Learned Mixture Model (AMM) in one of a supervised, semi supervised or unsupervised training.

5 According to an embodiment, the indication of identifiable entities comprises one of a number of classes and an indication of a class corresponding to at least one of said data.

According to an embodiment, the indication of identifiable entities comprises at least one box locating at least one corresponding identifiable entity.

10 According to a broad aspect, there is disclosed a non-transitory computer readable storage medium for storing computer-executable instructions which, when executed, cause a computer to perform a method for generating synthetically anonymized data for a given task, the method comprising providing first data to be anonymized; providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the  
15 first data; providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data; providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enables a disentanglement of different classes relevant to the given task; generating synthetically anonymized data for the  
20 given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to  
25 the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data; and providing the generated synthetically anonymized data for the given task.

According to another broad aspect, there is disclosed a computer comprising a central processing unit; a display device; a communication unit; a memory unit comprising an application for generating synthetically anonymized data for a given task, the application comprising instructions for providing first data to be anonymized; instructions for providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data; instructions for providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data; instructions for providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enables a disentanglement of different classes relevant to the given task; instructions for generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data; and instructions for providing the generated synthetically anonymized data for the given task.

It is an object to provide a method and a system which by design ensure anonymization of data based on an amendment of a defined set of identifiable features in data to prevent a re-identifying of the data.

It is another object to provide a method and a system which by design ensure that synthetic anonymized data conveys a suitable representation for processing the anonymized data for a given task.

The method disclosed herein is of great advantage for various reasons.

In fact, a first advantage of the method disclosed is that it provides privacy by-design for an anonymization process, while ensuring that the anonymized data is relevant for further research pertaining to a given task and to be representative of the general “look’n’feel” of the original data.

- 5 A second advantage of the method disclosed herein is that it enables the sharing of patient data in an open innovation environment, while ensuring patient privacy and control over the specific characteristics of the anonymized data (representative of all patient or sub-population thereof, representative globally of a task or sub-classes thereof).
- 10 A third advantage of the method disclosed herein is that it provides ways to anonymize data without an a-priori on what aspects of the data may convey such privacy risk(s); accordingly as such risk evolves, the method disclosed herein may adapt and benefit from further research and development in the field of data privacy.

15 **BRIEF DESCRIPTION OF THE DRAWINGS**

In order that the invention may be readily understood, embodiments of the invention are illustrated by way of example in the accompanying drawings.

- Figure 1 is a flowchart which shows an embodiment of a method for generating synthetically anonymized data for a given task. The method comprises *inter alia*,  
20 providing a task-specific embedding comprising task-specific features. The method further comprises providing an identifier embedding comprising identifiable features.

Figure 2 is a flowchart which shows an embodiment for providing an identifier embedding comprising identifiable features.

- Figure 3 is a flowchart which shows an embodiment for providing the task-specific  
25 embedding comprising task-specific features.

Figure 4 is a diagram which shows an embodiment of a system for generating synthetically anonymized data for a given task.

Figure 5 is a diagram which shows an embodiment of an Adversarially Learned Mixture Model (AMM) which may be used in an embodiment of the method for  
5 generating synthetically anonymized data for a given task.

Further details of the invention and its advantages will be apparent from the detailed description included below.

### **DETAILED DESCRIPTION**

10 In the following description of the embodiments, references to the accompanying drawings are by way of illustration of an example by which the invention may be practiced.

#### **Terms**

The term "invention" and the like mean "the one or more inventions disclosed in this application," unless expressly specified otherwise.

15 The terms "an aspect," "an embodiment," "embodiment," "embodiments," "the embodiment," "the embodiments," "one or more embodiments," "some embodiments," "certain embodiments," "one embodiment," "another embodiment" and the like mean "one or more (but not all) embodiments of the disclosed invention(s)," unless expressly specified otherwise.

20 A reference to "another embodiment" or "another aspect" in describing an embodiment does not imply that the referenced embodiment is mutually exclusive with another embodiment (e.g., an embodiment described before the referenced embodiment), unless expressly specified otherwise.

25 The terms "including," "comprising" and variations thereof mean "including but not limited to," unless expressly specified otherwise.

The terms "a," "an" and "the" mean "one or more," unless expressly specified otherwise.

The term "plurality" means "two or more," unless expressly specified otherwise.

5 The term "herein" means "in the present application, including anything which may be incorporated by reference," unless expressly specified otherwise.

10 The term "whereby" is used herein only to precede a clause or other set of words that express only the intended result, objective or consequence of something that is previously and explicitly recited. Thus, when the term "whereby" is used in a claim, the clause or other words that the term "whereby" modifies do not establish specific further limitations of the claim or otherwise restricts the meaning or scope of the claim.

The term "e.g." and like terms mean "for example," and thus do not limit the terms or phrases they explain.

15 The term "i.e." and like terms mean "that is," and thus limit the terms or phrases they explain.

20 The term "disentanglement" and like terms means in the real world that a models seek to represent, there are some factors of variation that can be modified independently, and others that cannot be (or, for practical purposes, never are). A trivial example of this is: if you're modeling pictures of people, then someone's clothing is independent of their height, whereas the length of their left leg is strongly dependent on the length of their right leg. The goal of disentangled features can be most easily understood as wanting to use each dimension of a latent z code to encode one and only one of these underlying independent factors of variation. Using the example from above, a disentangled representation would represent someone's height and clothing as separate dimensions of the z code.

25

The term “embedding” and like terms means relatively low-dimensional space into which high-dimensional vectors (dimensionality reduction) can be translated into. Embeddings make it easier to do machine learning on large inputs such as sparse vectors representing words or image characteristics. Ideally, an embedding captures  
5 some of the semantics of the input by placing semantically similar inputs close together (contextual similarity) in the embedding space. It will be appreciated that an embedding can be learned and reused across models. The purpose of an embedding is to map any input object (e.g. word, image) into vectors of real numbers, which algorithms, like deep learning, can then ingest and process, to formulate an  
10 understanding. The individual dimensions in these vectors typically have no inherent meaning. Instead, it is the overall patterns of location and distance between vectors that machine learning takes advantage of.

The term “feature” and like terms means, in machine learning and pattern recognition, an individual measurable property or characteristic of a phenomenon being observed.  
15 The concept of “feature” is related to that of explanatory variable used in statistical techniques such as linear regression. A feature vector is an n-dimensional vector of numerical features that represent some object. The vector space associated with these vectors is often called the feature space. In machine learning, feature learning or representation learning is a set of techniques that enables a system to  
20 automatically discover the representations needed for feature detection or classification from raw data. This replaces manual feature engineering and allows a machine to both learn the features and use them to perform a specific task. A classifier or neural network needs to be trained to learn to extract features from data. The features learned by a neural network depend among other things on the cost  
25 function used during training. The cost function defines the task that has to be solved. In order to have the ability to classify, the network is trained to minimize the classification error over training points. The embedding encodes the features extracted from the data. Multilayer neural networks can be used to perform feature learning, since they learn a representation of their input at the hidden layer(s) which is

subsequently used for classification or regression at the output layer. Deep neural networks learn feature embeddings of the input data that enable state-of-the-art performance in a wide range of computer vision tasks.

The term “generative” and like terms means a way of learning any kind of data distribution using unsupervised learning and it has achieved tremendous success in just a few years. All types of generative models aim at learning the true data distribution of the training set so as to generate new data points with some variations. But it is not always possible to learn the exact distribution of the data either implicitly or explicitly and so we try to model a distribution which is as similar as possible to the true data distribution. Two of the most commonly used and efficient approaches are Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN). Variational Autoencoders (VAE) aim at maximizing the lower bound of the data log-likelihood and Generative Adversarial Networks (GAN) aim at achieving an equilibrium between generator and discriminator.

Sampling - in Generative modeling with sampling can be considered one of the hardest tasks, it implies the ability to generate data that resemble the data used during training in the sense that they should ideally follow the same, unknown, true distribution. If data  $x$  are generated from an unknown distribution  $p$  such that  $x \sim p(x)$   $p$  can be approximated by learning a distribution  $q$ , from which it is possible to efficiently sample, that is close enough to  $p$ . This task is intimately related to probabilistic modeling and probability density estimation, but the focus is on the ability to generate good samples efficiently, rather than obtaining a precise numerical estimation of the probability density at a given point. There is a direct relation between “Generative” since sampling can generate synthetic data points.

Neither the Title nor the Abstract is to be taken as limiting in any way as the scope of the disclosed invention(s). The title of the present application and headings of sections provided in the present application are for convenience only, and are not to be taken as limiting the disclosure in any way.

Numerous embodiments are described in the present application, and are presented for illustrative purposes only. The described embodiments are not, and are not intended to be, limiting in any sense. The presently disclosed invention(s) are widely applicable to numerous embodiments, as is readily apparent from the disclosure.

5 One of ordinary skill in the art will recognize that the disclosed invention(s) may be practiced with various modifications and alterations, such as structural and logical modifications. Although particular features of the disclosed invention(s) may be described with reference to one or more particular embodiments and/or drawings, it should be understood that such features are not limited to usage in the one or more  
10 particular embodiments or drawings with reference to which they are described, unless expressly specified otherwise.

With all this in mind, the present invention is directed to a method and a system for generating synthetically anonymized data for a given task.

It will be appreciated that the method may be used in various embodiments. For  
15 instance in the medical field, the method may be used for generating synthetically anonymized patient data.

It will be appreciated that the given task to perform may be of various types.

In fact, the given task to perform is defined as any task in which the data may be used to.

20 For instance, in the medical field, the given task to perform may be used in one embodiment to determine an outcome of a patient in response to a treatment. In one embodiment, the given task to perform may be to provide a diagnostic. In another embodiment, the given task to perform may be one of anomaly detection and location (e.g. on images, on 1-D longitudinal information such as EKG), precision medicine  
25 prediction from various input information (e.g. images, clinical reports, EHR patient history), treatment strategy clinical decision support, drug side-effect prediction, relapse and metastasis prediction, readmission rate, post-operative surgical

complication, assisted surgery and assisted robotic surgery, preventative health prediction (e.g. Alzheimer, Parkinson, cardiac event or depression predictions).

It will be appreciated that the method and the system disclosed are of great advantage for many reasons, as explained further below.

- 5 Now referring to Fig. 1, there is shown an embodiment of a method for generating synthetically anonymized data for a given task.

It will be appreciated that the data may be any type of data which may be identified.

- For instance and in accordance with an embodiment, the data comprises patient data. The skilled addressee will appreciate that the patient data may be identifiable since it  
10 is associated with a given patient.

In another embodiment, the data is one of patient image data (e.g. CT scans, MRI, ultrasound, PET, X-rays), clinical reports, lab and pharmacy reports.

- It will be appreciated that the task is a processing to be performed using the data, to further predict downstream aspects related to the data, or classify the data. Generally  
15 speaking, a task may refer to one of a regression, a classification, a clustering, a multivariate querying, a density estimation, a dimension reduction and a testing and matching.

- It will be appreciated that the method disclosed herein for generating synthetically anonymized data for a given task may be implemented according to various  
20 embodiments.

Now referring to Fig. 4, there is shown an embodiment of a system for implementing the method disclosed herein for generating synthetically anonymized data for a given task. In this embodiment, the system comprises a computer 400. It will be appreciated that the computer 400 may be any type of computer.

In one embodiment, the computer 400 is selected from a group consisting of desktop computers, laptop computers, tablet PC's, servers, smartphones, etc. It will also be appreciated that, in the foregoing, the computer 400 may also be broadly referred to as a processor.

- 5 In the embodiment shown in Fig. 4, the computer 400 comprises a central processing unit (CPU) 402, also referred to as a microprocessor, input/output devices 404, a display device 406, a communication unit 408, a data bus 410 and a memory unit 412.

The central processing unit 402 is used for processing computer instructions. The skilled addressee will appreciate that various embodiments of the central processing unit 402 may be provided.

In one embodiment, the central processing unit 402 comprises a CPU Core i5 3210 running at 2.5 GHz and manufactured by Intel<sup>(TM)</sup>.

The input/output devices 404 are used for inputting/outputting data into the computer 400.

The display device 406 is used for displaying data to a user. The skilled addressee will appreciate that various types of display device 406 may be used.

In one embodiment, the display device 406 is a standard liquid crystal display (LCD) monitor.

20 The communication unit 408 is used for sharing data with the computer 400.

The communication unit 408 may comprise, for instance, universal serial bus (USB) ports for connecting a keyboard and a mouse to the computer 400.

The communication unit 408 may further comprise a data network communication port such as an IEEE 802.3 port for enabling a connection of the computer 400 with a remote processing unit, not shown.

The skilled addressee will appreciate that various alternative embodiments of the communication unit 408 may be provided.

The memory unit 412 is used for storing computer-executable instructions.

5 The memory unit 412 may comprise a system memory such as a high-speed random access memory (RAM) for storing system control program (e.g., BIOS, operating system module, applications, etc.) and a read-only memory (ROM).

It will be appreciated that the memory unit 412 comprises, in one embodiment, an operating system module 414.

It will be appreciated that the operating system module 414 may be of various types.

10 In one embodiment, the operating system module 414 is OS X Yosemite manufactured by Apple™. In another embodiment, the operating system module 414 comprises Linux Ubuntu 18.04.

The memory unit 412 further comprises an application for generating synthetically anonymized data 416.

15 The memory unit 412 further comprises models used by the application for generating synthetically anonymized data 416.

The memory unit 412 further comprises data used by the application for generating synthetically anonymized data 416.

20 Now referring back to Fig. 1 and according to processing step 100, a first data to be anonymized is provided.

It will be appreciated that the first data to be anonymized may be provided according to various embodiments. In accordance with an embodiment, the first data to be anonymized is obtained from the memory unit 412 of the computer 400.

In accordance with another embodiment, the first data to be anonymized is provided by a user interacting with the computer 400.

In accordance with yet another embodiment, the first data to be anonymized is obtained from a remote processing unit operatively coupled with the computer 400. It will be appreciated that the remote processing unit may be operatively coupled with the computer 400 according to various embodiments. In one embodiment, the remote processing unit is operatively coupled with the computer 400 via a data network selected from a group comprising at least one of a local area network, a metropolitan area network and a wide area network. In one embodiment, the data network comprises the Internet.

As mentioned above, it will be appreciated that in one embodiment the first data to be anonymized comprises patient data.

According to processing step 101, a data embedding comprising data features is provided. It will be appreciated that the data features enable a representation of corresponding data and the data is representative of the first data.

In one embodiment, the data embedding is obtained by training a deep generative model in a representation learning task, onto the data itself, such as disclosed in “representation learning: a review and new perspectives - arXiv:1206.5538”, in “Variational lossy autoencoder. arXiv:1611.02731”, in “neural discrete representation learning - arXiv:1711.00937” and in “Privacy-preserving generative deep neural networks support clinical data sharing - bioarxiv:159756”.

Moreover, it will be appreciated that the data embedding may be provided according to various embodiments. In accordance with an embodiment, the data embedding is obtained from the memory unit 412 of the computer 400.

In accordance with another embodiment, the data embedding is provided by a user interacting with the computer 400.

In accordance with yet another embodiment, the data embedding is obtained from a remote processing unit operatively coupled with the computer 400.

Still referring to Fig. 1 and according to processing step 102, an identifier embedding comprising identifiable features is provided. It will be appreciated that the identifiable features enable an identification of the data and the first data.

It will be appreciated by the skilled addressee that the identifier embedding comprising identifiable features may be provided according to various embodiments.

Now referring to Fig. 2, there is shown an embodiment for providing the identifier embedding comprising the identifiable features.

10 According to processing step 200, data used for identifying the identifiable features is obtained.

It will be appreciated that the data used for identifying features may be of various types. In one embodiment, the data used for identifying the identifiable features comprises at least one portion of the first data provided.

15 In accordance with another embodiment, the data used for identifying the identifiable features may be different data than the first data provided according to processing step 100.

It will be also appreciated that the data used for identifying the identifiable features may be provided according to various embodiments.

20 In accordance with an embodiment, the data used for identifying the identifiable features is obtained from the memory unit 412 of the computer 400.

In accordance with another embodiment, the data used for identifying the identifiable features is provided by a user interacting with the computer 400.

In accordance with yet another embodiment, the data used for identifying the identifiable features is obtained from a remote processing unit operatively coupled with the computer 400, as explained above.

5 According to processing step 202, a model suitable for identifying the identifiable features is obtained.

In one embodiment, the model suitable for identifying the identifiable features is a Single Shot MultiBox Detector (SSD) model known to the skilled addressee. The skilled addressee will appreciate that various alternative embodiments may be provided for the model suitable for identifying the identifiable features. For instance  
10 and in accordance with another embodiment, the model suitable for identifying the identifiable features is a You Only Look Once (YOLO) model, known to the skilled addressee.

It will be also appreciated that the model suitable for identifying the identifiable features may be provided according to various embodiments.

15 In accordance with an embodiment, the model suitable for identifying the identifiable features is obtained from the memory unit 412 of the computer 400.

In accordance with another embodiment, the model suitable for identifying the identifiable features is provided by a user interacting with the computer 400.

In accordance with yet another embodiment, the model suitable for identifying the  
20 identifiable features is obtained from a remote processing unit operatively coupled with the computer 400 as explained above.

Still referring to Fig. 2 and according to processing step 204, an indication of identifiable entities is provided.

It will be appreciated that the indication of identifiable entities refers to elements that  
25 may be used to identify data such as morphometric patterns in imaging data, acoustic pattern in spectral data (albeit spectrogram), trending pattern in 1-D data.

For instance and in the case of patient data, the identifiable entities refer to elements that may be used to identify a patient.

In the context of imaging patient data, organs could be used to identify patient data, and accordingly said indication of identifiable entities could be a weak indication of organs' presence at the level of imaging patient data, organ bounding boxes on some imaging patient data, organ segmentation on some imaging patient data. Additional elements that may be used to identify patients are morphometry of the face either directly or indirectly obtained in the case of CT of the head for example, gait from videos, patient history and chronology of specific events, patient-specific morphology either from birth defects or surgically related.

It will be also appreciated that the indication of identifiable entities may be provided according to various embodiments.

In accordance with an embodiment, the indication of identifiable entities is obtained from the memory unit 412 of the computer 400.

In accordance with another embodiment, the indication of identifiable entities is provided by a user interacting with the computer 400.

In accordance with yet another embodiment, the indication of identifiable entities is obtained from a remote processing unit operatively coupled with the computer 400 as explained above.

Still referring to Fig. 2 and according to processing step 206, an identifier embedding is generated.

It will be appreciated that the identifier embedding is generated using the model suitable for identifying the identifiable features, the indication of identifiable entities and the data to be used for identifying the identifiable features.

In one embodiment, the identifier embedding is generated using the computer 400.

Now referring back to Fig. 1 and according to processing step 104, a task-specific embedding comprising task-specific features is generated.

It will be appreciated that the task-specific embedding comprising task-specific features may be generated according to various embodiments.

- 5 Now referring to Fig. 3, there is shown an embodiment for generating the task-specific embedding comprising task-specific features.

According to processing step 300, an indication of the given task is obtained.

As mentioned above, it will be appreciated that the indication of the given task may be of various types.

- 10 It will be also appreciated that the indication of the given task may be provided according to various embodiments.

In accordance with an embodiment, the indication of the given task is obtained from the memory unit 512 of the computer 500.

- 15 In accordance with another embodiment, the indication of the given task is provided by a user interacting with the computer 500.

In accordance with yet another embodiment, the indication of the given task is obtained from a remote processing unit operatively coupled with the computer 500 as explained above.

- 20 Still referring to Fig. 3 and according to processing step 302, an indication of classes relevant to the given task is provided.

It will be appreciated by the skilled addressee that the indication of classes relevant to the given task are at least binary, for instance responding, nonresponding - malignant/benign, or multi-classes, such as for instance disease progression, no progression, pseudo-progression.

It will be also appreciated that the indication of classes relevant to the given task may be provided according to various embodiments.

In accordance with an embodiment, the indication of classes relevant to the given task is obtained from the memory unit 412 of the computer 400.

- 5 In accordance with another embodiment, the indication of classes relevant to the given task is provided by a user interacting with the computer 400.

In accordance with yet another embodiment, the indication of classes relevant to the given task is obtained from a remote processing unit operatively coupled with the computer 400 as explained above.

- 10 Still referring to Fig. 3 and according to processing step 304, a model suitable for performing a disentanglement of the first data is provided.

In one embodiment, the model suitable for performing a disentanglement of the first data is the Adversarially Learned Mixture Model (AMM) disclosed herein.

- 15 It will be appreciated that alternative embodiments of the model suitable for performing a disentanglement of the data may be provided. In fact, it has been contemplated that any model capable of modeling complex data distribution may be used. It will be appreciated that the Generative Adversarial Network (GAN) has recently emerged as a powerful framework for modeling complex data distributions without having to approximate intractable likelihoods. As mentioned above and in a preferred embodiment an Adversarially Learned Mixture Model (AMM) is used, a  
20 generative model inferring both continuous and categorical latent variables to perform either unsupervised or semi-supervised clustering of data using a single adversarial objective, that explicitly model the dependence between continuous and categorical latent variables, and which eliminates discontinuities between categories in the latent  
25 space.

It will be also appreciated that the model suitable for performing a disentanglement of the first data may be provided according to various embodiments.

5 In accordance with an embodiment, the model suitable for performing a disentanglement of the first data is obtained from the memory unit 412 of the computer 400.

In accordance with another embodiment, the model suitable for performing a disentanglement of the first data is provided by a user interacting with the computer 400.

10 In accordance with yet another embodiment, the model suitable for performing a disentanglement of the first data is obtained from a remote processing unit operatively coupled with the computer 400 as explained above.

Still referring to Fig. 3 and according to processing step 306, a task-specific embedding is generated.

15 It will be appreciated that a task-specific embedding refers to one of a regression, a classification, a clustering, a multivariate querying, a density estimation, a dimension reduction and a testing and matching.

More precisely, the task-specific embedding is generated using the obtained model, the indication of classes relevant to the given task, the indication of the given task and the data. In another embodiment, the task-specific embedding is generated using the  
20 obtained model, the indication of classes relevant to the given task, the indication of the given task and the first data.

Such generation of the task-embedding can be performed, in a preferred embodiment, using the above-mentioned Adversarially Learned Mixture Model (AMM). In another embodiment, a generative model following “Learning disentangled  
25 representations with semi-supervised deep generative models - arXiv:1706.00400 [stat.ML]” may be used

Now referring back to Fig. 1 and according to processing step 106, the synthetically anonymized data for the given task is generated.

It will be appreciated that the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a  
5 corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features. The generating further mixes the first sample and the second sample in a generative process to create the generated synthetically  
10 anonymized data.

In one embodiment, the first sampling from the data embedding which ensures that corresponding first sample originates away from a projection of the data and the first data in the identifier embedding is performed using a rejection sampling technique such as detailed in "Deep Learning for Sampling from Arbitrary Probability  
15 Distributions - arXiv:1801.04211".

In another embodiment, the sampling process is performed using a Markov chain Monte Carlo (MCMC) sampling process such as detailed in "Improving Sampling from GenerativeAutoencoders with Markov Chains - OpenReview ryXZmzNeg - Antonia Creswell, Kai Arulkumaran, Anil Anthony Bharath 30 Oct 2016 (modified: 12 Jan  
20 2017) ICLR 2017 conference submission"; accordingly, since, the generative model learns to map from the learned latent distribution, rather than the prior, a Markov chain Monte Carlo (MCMC) sampling process may be used to improve the quality of samples drawn from the generative model, especially when the learned latent distribution is far from the prior.

25 In yet a further embodiment, the sampling process includes Parallel Checkpointing Learners methods that ensure that although samples originates away from a projected a-priori known data in the identifiable embedding, the generative model is robust against adversarial samples, by rejecting samples that are likely to come from

the unexplored regions conveying potentially high risk of irrelevance such as detailed in "Towards Safe Deep Learning: Unsupervised Defense Against Generic Adversarial Attacks - OpenReview Hyl6s40a-".

In one embodiment, mixing samples originating from different embeddings is performed as disclosed in "conditional generative adversarial nets - arXiv:1411.1784", in "Generative adversarial text to image synthesis - arXiv:1605.05396", in "PixelBrush: Art generation from text with GANs - Jiale Zhi Stanford University" and in "RenderGAN: generating realistic labelled data - arXiv:1611.01331".

Still referring to Fig. 1 and according to processing step 108, a check is performed in order to find out if the generated synthetically anonymized data is dissimilar to the first data to be anonymized for a given metric. It will be appreciated that processing step 108 is optional.

It will be appreciated that the given metric may be of various types as known to the skilled addressee.

In fact and in one embodiment, the checking that the generated synthetically anonymized data is dissimilar to the first data to be anonymized for a given metric, is performed following traditional image similarity measures as detailed in "Mitchell H.B. (2010) Image Similarity Measures. In: Image Fusion. Springer, Berlin, Heidelberg", or following differential privacy as detailed in "Privacy-preserving generative deep neural networks support clinical data sharing - bioarxiv:159756", in "L. Sweeney, k-anonymity: A model for protecting privacy, Int. J. Uncertainty, Fuzziness (2002)".

While it has been disclosed that the checking is performed following the generating step 106, it will be appreciated by the skilled addressee that in another alternative embodiment, the checking performed according to processing step 108 is integrated in the generating processing step disclosed in processing step 106 as detailed in "Generating differentially private datasets using GANs - OpenReview rJv4XWZA-, ICLR 2018". In such embodiment, the checking step as disclosed in Fig. 1 is optional.

In such embodiment, the generating of the synthetically anonymized data for the given task comprises checking that the synthetically anonymized data is dissimilar to the first data to be anonymized for a given metric.

5 According to processing step 110, the generated synthetically anonymized data for the given task is provided. It will be appreciated that the generated synthetically anonymized data for the given task is provided if the checking is successful.

It will be appreciated that the generated synthetically anonymized data may be provided according to various embodiments.

10 In accordance with an embodiment, the generated synthetically anonymized data is stored in the memory unit 412 of the computer 400.

In accordance with another embodiment, the generated synthetically anonymized data is provided to a remote processing unit operatively coupled to the computer 400.

In another alternative embodiment, the generated synthetically anonymized data is displayed to a user interacting with the computer 400.

15 Still referring to Fig. 4, it will be appreciated that the application for generating synthetically anonymized data 416 comprises instructions for providing first data to be anonymized.

20 The application for generating synthetically anonymized data 416 further comprises instructions for providing a data embedding comprising data features, wherein data features enable a representation of corresponding data wherein the data is representative of the first data.

25 The application for generating synthetically anonymized data 416 further comprises instructions for providing an identifier embedding comprising identifiable features. It will be appreciated that the identifiable features enable an identification of the first data.

The application for generating synthetically anonymized data 416 further comprises instructions for providing a task-specific embedding comprising task specific features suitable for the task. It will be appreciated that the task specific features enable a disentanglement of different classes relevant to the given task.

- 5 The application for generating synthetically anonymized data for the given task further comprises instructions for generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projected data and the first data in  
10 the identifiable embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data.
- 15 The application for generating synthetically anonymized data for the given task further comprises instructions for checking that the synthetically anonymized data is dissimilar to the first data to be anonymized for a given metric.

The application for generating synthetically anonymized data for the given task further comprises instructions for providing the generated synthetically anonymized data for  
20 the given task if said checking is successful.

A non-transitory computer readable storage medium is disclosed for storing computer-executable instructions which, when executed, cause a computer to perform a method for generating synthetically anonymized data for a given task, the method comprising providing first data to be anonymized; providing a data embedding  
25 comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data; providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data; providing a task-specific embedding

comprising task-specific features suitable for said task, wherein said task-specific features enables a disentanglement of different classes relevant to the given task; generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projected data and the first data in the identifiable embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data; checking that the synthetically anonymized data is dissimilar to the first data to be anonymized for a given metric and providing the generated synthetically anonymized data for the given task if the checking is successful.

It will be appreciated that the method disclosed herein is of great advantage for various reasons.

In fact, a first advantage of the method disclosed is that it provides privacy by-design for an anonymization process, while ensuring that the anonymized data is relevant for further research pertaining to a given task and to be representative of the general “look’n’feel” of the original data.

A second advantage of the method disclosed herein is that it enables the sharing of patient data in an open innovation environment, while ensuring patient privacy and control over the specific characteristics of the anonymized data (representative of all patient or sub-population thereof, representative globally of a task or sub-classes thereof).

A third advantage of the method disclosed herein is that it provides ways to anonymize data without a-priori on what aspects of the data may convey such privacy risk(s); accordingly as such risk evolve, the method disclosed herein may adapt and benefit from further research and development in the field of data privacy.

### Adversarially Learned Mixture Model (AMM)

It will be appreciated that the Adversarially Learned Mixture Model (AMM) is disclosed herein below. This model may be used advantageously in the method disclosed herein as mentioned previously.

- 5 It is known to the skilled addressee that the ALI and BiGAN models are trained by matching two joint distributions of images  $\mathbf{x} \in \mathbb{R}^D$  and their latent code  $\mathbf{z} \in \mathbb{R}^L$ . The two distributions to be matched are the inference distribution  $q(\mathbf{x}, \mathbf{z})$  and the synthesis distribution  $p(\mathbf{x}, \mathbf{z})$ , wherein,

$$q(\mathbf{x}, \mathbf{z}) = q(\mathbf{x})q(\mathbf{z} | \mathbf{x}), \quad \text{Equation (1)}$$

10  $p(\mathbf{x}, \mathbf{z}) = p(\mathbf{z})p(\mathbf{x} | \mathbf{z}). \quad \text{Equation (2)}$

Samples of  $q(\mathbf{x})$  are drawn from the training data and samples of  $p(\mathbf{z})$  are drawn from a prior distribution, usually  $\mathcal{N}(0,1)$ . Samples from  $q(\mathbf{z} | \mathbf{x})$  and  $p(\mathbf{x} | \mathbf{z})$  are drawn from neural networks that are optimized during training. Dumoulin et al. (See “Adversarially learned inference”. in *International Conference on Learning Representation* (2016))

- 15 show that sampling from  $q(\mathbf{z} | \mathbf{x}) = \mathcal{N}(\mu(\mathbf{x}), \sigma^2(\mathbf{x})I)$  is possible by employing the reparametrization trick (See Kingma & Welling, “Auto-encoding variational Bayes”, in *International Conference on Learning Representation* (2013)), i.e. computing:

$$\mathbf{z} = \mu(\mathbf{x}) + \sigma(\mathbf{x}) \odot \epsilon, \epsilon \sim \mathcal{N}(0, I), \quad \text{Equation (3)}$$

wherein  $\odot$  is the element wise vector multiplication.

- 20 A conditional variant of ALI has also been explored by Dumoulin et al. (2016) wherein an observed class-conditional categorical variable  $\mathbf{y}$  has been introduced. The joint factorization of each distribution to be matched are:

$$q(\mathbf{x}, \mathbf{y}, \mathbf{z}) = q(\mathbf{x}, \mathbf{y})q(\mathbf{z} | \mathbf{y}, \mathbf{x}), \quad \text{Equation (4)}$$

$$p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = p(\mathbf{y})p(\mathbf{z})q(\mathbf{x} | \mathbf{y}, \mathbf{z}). \quad \text{Equation (5)}$$

It will be appreciated that samples of  $q(\mathbf{x}, \mathbf{y})$  are drawn from the data, samples of  $p(\mathbf{z})$  are drawn from a continuous prior on  $\mathbf{z}$ , and samples of  $p(\mathbf{y})$  are drawn from a categorical prior on  $\mathbf{y}$ , both of which are marginally independent. It will be further appreciated that samples from  $q(\mathbf{z} | \mathbf{y}, \mathbf{x})$  and  $p(\mathbf{x} | \mathbf{y}, \mathbf{z})$  are drawn from neural  
 5 networks that are optimized during training.

In the following, graphical models are presented for  $q(\mathbf{x}, \mathbf{y}, \mathbf{z})$  and  $p(\mathbf{x}, \mathbf{y}, \mathbf{z})$  that build off of conditional ALI. Where conditional ALI requires the full observation of categorical variables, the models presented account for both unobserved and partially observed categorical variables.

#### 10 *Adversarially learned mixture model*

It will be appreciated that the Adversarially Learned Mixture Model (AMM) disclosed herein and illustrated in Fig. 5 is an adversarial generative model for deep unsupervised clustering of data.

Like conditional ALI, a categorical variable is introduced to model the labels.

15 However, the unsupervised setting requires a different factorization of the inference distribution in order to enable inference of the categorical variable  $\mathbf{y}$ , namely:

$$q_1(\mathbf{x}, \mathbf{y}, \mathbf{z}) = q(\mathbf{x})q(\mathbf{y} | \mathbf{x})q(\mathbf{z} | \mathbf{x}, \mathbf{y}), \quad \text{Equation (6)}$$

or

$$q_2(\mathbf{x}, \mathbf{y}, \mathbf{z}) = q(\mathbf{x})q(\mathbf{z} | \mathbf{x})q(\mathbf{y} | \mathbf{x}, \mathbf{z}). \quad \text{Equation (7)}$$

20 Samples of  $q(\mathbf{x})$  are drawn from the training data, and samples from  $q(\mathbf{y} | \mathbf{x})$ ,  $q(\mathbf{z} | \mathbf{x}, \mathbf{y})$  or  $q(\mathbf{z} | \mathbf{x})$ ,  $q(\mathbf{y} | \mathbf{x}, \mathbf{z})$  are generated by neural networks. It will be appreciated that the reparametrization trick is not directly applicable to discrete variables and multiple methodologies have been introduced to approximate categorical samples (See Jang et al. "Categorical reparametrization with Gumbel-softmax". arXiv preprint  
 25 arXiv:1611.01144, 2016; Maddison et al. "The concrete Distribution: A Continuous

Relaxation of Discrete Random Variables.” in *International Conference on learning representations*, 2017). It will be appreciated that in this embodiment Kendall & Gal (See “What uncertainties do we need in Bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems* 30, pp. 5580-5590 (2017)) is followed and a sample is performed from  $q(\mathbf{y} | \mathbf{x})$  by computing:

$$h_{\mathbf{y}}(\mathbf{x}) = \mu(\mathbf{x}) + \sigma_{\mathbf{y}}(\mathbf{x}) \odot \epsilon, \epsilon \sim \mathcal{N}(0, I), \quad \text{Equation (8)}$$

$$\mathbf{y}(\mathbf{x}) = \text{softmax}(h_{\mathbf{y}}(\mathbf{x})). \quad \text{Equation (9)}$$

It is then possible to sample from  $q(\mathbf{z} | \mathbf{x}, \mathbf{y})$ , by computing:

$$\mathbf{z}(\mathbf{x}, h_{\mathbf{y}}(\mathbf{x})) = \mu_{\mathbf{z}}(\mathbf{x}, h_{\mathbf{y}}(\mathbf{x})) + \sigma_{\mathbf{z}}(\mathbf{x}, h_{\mathbf{y}}(\mathbf{x})) \odot \epsilon, \epsilon \sim \mathcal{N}(0, I). \quad \text{Equation (10)}$$

10 A similar sampling strategy may be used to sample from  $q(\mathbf{y} | \mathbf{x}, \mathbf{z})$  in Equation (7).

The factorization of the synthesis distribution  $p(\mathbf{x}, \mathbf{y}, \mathbf{z})$  also differs from conditional ALI:

$$p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = p(\mathbf{y})p(\mathbf{z} | \mathbf{y})p(\mathbf{x} | \mathbf{y}, \mathbf{z}). \quad \text{Equation (11)}$$

15 It will be appreciated that the product  $p(\mathbf{y})p(\mathbf{z} | \mathbf{y})$  may be conveniently given by a mixture model. Samples from  $p(\mathbf{y})$  are drawn from a multinomial prior, and samples from  $p(\mathbf{z} | \mathbf{y})$  are drawn from a continuous prior, for example,  $\mathcal{N}(\mu_{\mathbf{y}=k}, 1)$ . Samples from  $p(\mathbf{z} | \mathbf{y})$  may alternatively be generated by a neural network by again employing the reparameterization trick, namely:

$$\mathbf{z}(\mathbf{y}) = \mu(\mathbf{y}) + \sigma(\mathbf{y}) \odot \epsilon, \epsilon \sim \mathcal{N}(0, I). \quad \text{Equation (12)}$$

20 This approach effectively learns the parameters of  $\mathcal{N}(\mu_{\mathbf{y}=k}, \sigma_{\mathbf{y}=k})$ .

#### *Adversarial value function*

Dumoulin et al. (2016) is followed and the value function that describes the unsupervised game between the discriminator  $D$  and the generator  $G$  is defined as:

$$\begin{aligned} \min_G \max_D V(D, G) = & \mathbb{E}_{q(x)} \left[ \log \left( D \left( x, G_y(x), G_z(x, G_y(x)) \right) \right) \right] + \mathbb{E}_{p(y, z)} \left[ \log \left( 1 - \right. \right. \\ & \left. \left. D \left( G_x(y, G_z(y)), y, G_z(y) \right) \right) \right] = \iiint q(x)q(y|x)q(z|x, y) \log(D(x, y, z)) dx dy dz + \\ & \iiint p(y)p(z|y)p(x|y, z) \log(1 - D(x, y, z)) dx dy dz \end{aligned} \quad \text{Equation (13)}$$

- 5 It will be appreciated that there are four generators in total: two for the encoder  $G_y(x)$  and  $G_z(x, G_y(x))$ , which map the data samples to the latent space; and two for the decoder  $G_z(y)$  and  $G_x(y, G_z(y))$ , which map samples from the prior to the input space.  $G_z(y)$  can either be a learned function, or be specified by a known prior. A detailed description of the optimization procedure is detailed herein below.

10

---

**Algorithm 1** AMM training procedure using distributions (6) and (11).

---

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\theta_{G_y(x)}, \theta_{G_z(x, G_y(x))}, \theta_{G_z(y)}, \theta_{G_x(y, G_z(y))}, \theta_D$<br><b>while</b> not done <b>do</b><br>$x^{(1)}, \dots, x^{(M)} \sim q(x)$<br>$y^{(1)}, \dots, y^{(M)} \sim p(y)$<br>$z^{(j)} \sim p(z   y = y^{(j)}), \quad j = 1, \dots, M$<br>$\tilde{x}^{(j)} \sim p(x   y = y^{(j)}, z = z^{(j)}), \quad j = 1, \dots, M$<br>$\tilde{y}^{(i)} \sim q(y   x = x^{(i)}), \quad i = 1, \dots, M$<br>$\tilde{z}^{(i)} \sim q(z   x = x^{(i)}, y = \tilde{y}^{(i)}), \quad i = 1, \dots, M$<br>$\rho_q^{(i)} \leftarrow D(x^{(i)}, \tilde{y}^{(i)}, \tilde{z}^{(i)}), \quad i = 1, \dots, M$<br>$\rho_p^{(j)} \leftarrow D(\tilde{x}^{(j)}, y^{(j)}, z^{(j)}), \quad j = 1, \dots, M$<br>$\mathcal{L}_D \leftarrow -\frac{1}{M} \sum_{i=1}^M \log(\rho_q^{(i)}) - \frac{1}{M} \sum_{j=1}^M \log(1 - \rho_p^{(j)})$<br>$\mathcal{L}_{G_z(y, G_z(y))} = \mathcal{L}_{G_z(y)} \leftarrow -\frac{1}{M} \sum_{i=1}^M \log(\rho_q^{(i)})$<br>$\mathcal{L}_{G_y(x)} = \mathcal{L}_{G_z(x, G_y(x))} \leftarrow -\frac{1}{M} \sum_{i=1}^M \log(1 - \rho_q^{(i)})$<br>$\theta_D \leftarrow \theta_D - \nabla_{\theta_D} \mathcal{L}_D$<br>$\theta_{G_z(y, G_z(y))} \leftarrow \theta_{G_z(y, G_z(y))} - \nabla_{\theta_{G_z(y, G_z(y))}} \mathcal{L}_{G_z(y, G_z(y))}$<br>$\theta_{G_z(y)} \leftarrow \theta_{G_z(y)} - \nabla_{\theta_{G_z(y)}} \mathcal{L}_{G_z(y)}$<br>$\theta_{G_y(x)} \leftarrow \theta_{G_y(x)} - \nabla_{\theta_{G_y(x)}} \mathcal{L}_{G_y(x)}$<br>$\theta_{G_z(x, G_y(x))} \leftarrow \theta_{G_z(x, G_y(x))} - \nabla_{\theta_{G_z(x, G_y(x))}} \mathcal{L}_{G_z(x, G_y(x))}$ | <p>▷ Initialize AMM parameters</p> <p>▷ Sample from data and priors</p> <p>▷ Sample from conditionals</p> <p>▷ Compute discriminator predictions</p> <p>▷ Compute discriminator losses</p> <p>▷ Compute x generator losses</p> <p>▷ Compute y and z generator loss</p> <p>▷ Update discriminator parameters</p> <p>▷ Update generator parameters</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

---

*Semi-supervised adversarially learned mixture model*

The Semi-Supervised Adversarially Learned Mixture Model (SAMM) is an adversarial generative model for supervised or semi-supervised clustering and classification of data. The objective for training the Semi-Supervised Adversarially Learned Mixture Model involves two adversarial games to match pairs of joint distributions. The supervised game matches inference distribution (4) to synthesis distribution (11) and is described by the following value function:

$$\begin{aligned} \min_G \max_D V(D, G) = & \mathbb{E}_{q(\mathbf{x}, \mathbf{y})} [\log(D(\mathbf{x}, \mathbf{y}, G_z(\mathbf{x}, \mathbf{y})))] + \mathbb{E}_{p(\mathbf{y}, \mathbf{z})} \left[ \log \left( 1 - \right. \right. \\ & \left. \left. D \left( G_x(\mathbf{y}, G_z(\mathbf{y})), \mathbf{y}, G_z(\mathbf{y}) \right) \right) \right] = \iiint q(\mathbf{x}, \mathbf{y}) q(\mathbf{z} | \mathbf{x}, \mathbf{y}) \log(D(\mathbf{x}, \mathbf{y}, \mathbf{z})) d\mathbf{x} d\mathbf{y} d\mathbf{z} + \\ & \iiint p(\mathbf{y}) p(\mathbf{z} | \mathbf{y}) p(\mathbf{x} | \mathbf{y}, \mathbf{z}) \log(1 - D(\mathbf{x}, \mathbf{y}, \mathbf{z})) d\mathbf{x} d\mathbf{y} d\mathbf{z}. \end{aligned}$$

Clauses:

Clause 1. A method for generating synthetically anonymized data for a given task, the method comprising:

providing first data to be anonymized;

providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data;

providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data;

providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enable a disentanglement of different classes relevant to the given task;

generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates

away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a  
5 generative process to create the generated synthetically anonymized data; and

providing the generated synthetically anonymized data for the given task.

Clause 2. The method as claimed in clause 1, wherein the generating of the synthetically anonymized data for the given task comprises checking that the synthetically anonymized data is dissimilar to the first data to be anonymized for a  
10 given metric; further wherein the generated synthetically anonymized data for the given task is provided if said checking is successful.

Clause 3. The method as claimed in any one of clauses 1 to 2, wherein the first data comprises patient data.

Clause 4. The method as claimed in any one of clauses 1 to 3, wherein the providing  
15 of the task-specific embedding comprising task specific features suitable for said task comprises:

obtaining an indication of the given task;

obtaining an indication of classes relevant to the given task;

obtaining a model suitable for performing a disentanglement of the data for  
20 the given task; and

generating the task-specific embedding using the obtained model, the indication of classes relevant to the given task, the indication of the given task and the data.

Clause 5. The method as claimed in any one of clauses 1 to 4, wherein the providing  
25 of the identifier embedding comprising identifiable features comprises:

obtaining data used for identifying the identifiable features;

obtaining a model suitable for identifying the identifiable features in said data;

obtaining an indication of identifiable entities; and

- 5           generating the identifier embedding using the model suitable for identifying the identifiable features, the indication of identifiable entities and the data to be used for identifying the identifiable features.

Clause 6. The method as claimed in clause 5, wherein the data comprises the data used for identifying the identifiable features.

- 10       Clause 7. The method as claimed in clause 5, wherein the model suitable for identifying the identifiable features in said data comprises a Single Shot MultiBox Detector (SSD) model.

- Clause 8. The method as claimed in clause 4, wherein the model suitable for performing a disentanglement of the data for the given task comprises one of an  
15       Adversarially Learned Mixture Model (AMM) in one of a supervised, semi supervised or unsupervised training.

Clause 9. The method as claimed in clause 4, wherein the indication of identifiable entities comprises one of a number of classes and an indication of a class corresponding to at least one of said data.

- 20       Clause 10. The method as claimed in clause 5, wherein the indication of identifiable entities comprises at least one box locating at least one corresponding identifiable entity.

- Clause 11. A non-transitory computer readable storage medium for storing computer-executable instructions which, when executed, cause a computer to perform a  
25       method for generating synthetically anonymized data for a given task, the method

comprising providing first data to be anonymized; providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data; providing an identifier embedding comprising identifiable features, wherein the identifiable  
5 features enable an identification of the data and the first data; providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enables a disentanglement of different classes relevant to the given task; generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first  
10 sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a  
15 generative process to create the generated synthetically anonymized data; and providing the generated synthetically anonymized data for the given task.

Clause 12. A computer comprising:

a central processing unit;

a display device;

20 a communication unit;

a memory unit comprising an application for generating synthetically anonymized data for a given task, the application comprising:

instructions for providing first data to be anonymized;

instructions for providing a data embedding comprising data features,  
25 wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data;

instructions for providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data;

5 instructions for providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enables a disentanglement of different classes relevant to the given task;

instructions for generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a  
10 corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically  
15 anonymized data; and

instructions for providing the generated synthetically anonymized data for the given task.

Although the above description relates to a specific preferred embodiment as presently contemplated by the inventor, it will be understood that the invention in its  
20 broad aspect includes functional equivalents of the elements described herein.

## CLAIMS:

1. A method for generating synthetically anonymized data for a given task, the method comprising:

providing first data to be anonymized, wherein the first data includes at least one of: patient image data, patient clinical reports data, patient electronic health record (EHR) data, and patient electronic medical record (EMR) data;

providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data;

providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data;

providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enable a disentanglement of different classes relevant to the given task;

generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data; and

providing the generated synthetically anonymized data for the given task.

2. The method as claimed in claim 1, wherein the generating of the synthetically anonymized data for the given task comprises checking that the synthetically anonymized data is dissimilar to the first data to be anonymized for a given metric;

further wherein the generated synthetically anonymized data for the given task is provided if said checking is successful.

3. The method as claimed in any one of claims 1 to 2, wherein the providing of the task-specific embedding comprising task specific features suitable for said task comprises:

- obtaining an indication of the given task;

- obtaining an indication of classes relevant to the given task;

- obtaining a model suitable for performing a disentanglement of the data for the given task; and

- generating the task-specific embedding using the obtained model, the indication of classes relevant to the given task, the indication of the given task and the data.

4. The method as claimed in any one of claims 1 to 3, wherein the providing of the identifier embedding comprising identifiable features comprises:

- obtaining data used for identifying the identifiable features;

- obtaining a model suitable for identifying the identifiable features in said data;

- obtaining an indication of identifiable entities; and

- generating the identifier embedding using the model suitable for identifying the identifiable features, the indication of identifiable entities and the data to be used for identifying the identifiable features.

5. The method as claimed in claim 4, wherein the data comprises the data used for identifying the identifiable features.

6. The method as claimed in claim 4, wherein the model suitable for identifying the identifiable features in said data comprises a Single Shot MultiBox Detector (SSD) model.

7. The method as claimed in claim 3, wherein the model suitable for performing a disentanglement of the data for the given task comprises one of an Adversarially Learned Mixture Model (AMM) in one of a supervised, semi supervised or unsupervised training.
8. The method as claimed in claim 3, wherein the indication of identifiable entities comprises one of a number of classes and an indication of a class corresponding to at least one of said data.
9. The method as claimed in claim 4, wherein the indication of identifiable entities comprises at least one box locating at least one corresponding identifiable entity.
10. A non-transitory computer readable storage medium for storing computer-executable instructions which, when executed, cause a computer to perform a method for generating synthetically anonymized data for a given task, the method comprising providing first data to be anonymized, wherein the first data includes at least one of: patient image data, patient clinical reports data, patient electronic health record (EHR) data, and patient electronic medical record (EMR) data; providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data; providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data; providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enables a disentanglement of different classes relevant to the given task; generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates

close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data; and providing the generated synthetically anonymized data for the given task.

11. A computer comprising:

a central processing unit;

a display device;

a communication unit;

a memory unit comprising an application for generating synthetically anonymized data for a given task, the application comprising:

instructions for providing first data to be anonymized, wherein the first data includes at least one of: patient image data, patient clinical reports data, patient electronic health record (EHR) data, and patient electronic medical record (EMR) data;

instructions for providing a data embedding comprising data features, wherein data features enable a representation of corresponding data, and wherein the data is representative of the first data;

instructions for providing an identifier embedding comprising identifiable features, wherein the identifiable features enable an identification of the data and the first data;

instructions for providing a task-specific embedding comprising task-specific features suitable for said task, wherein said task-specific features enables a disentanglement of different classes relevant to the given task;

instructions for generating synthetically anonymized data for the given task, wherein the generating comprises a generative process using samples comprising a first sampling from the data embedding which ensures that a corresponding first sample originates away from a projection of the data and the first

data in the identifier embedding and a second sampling from the task-specific embedding which ensures that a corresponding second sample originates close to the task-specific features and wherein the generating further mixes the first sample and the second sample in a generative process to create the generated synthetically anonymized data; and

instructions for providing the generated synthetically anonymized data for the given task.

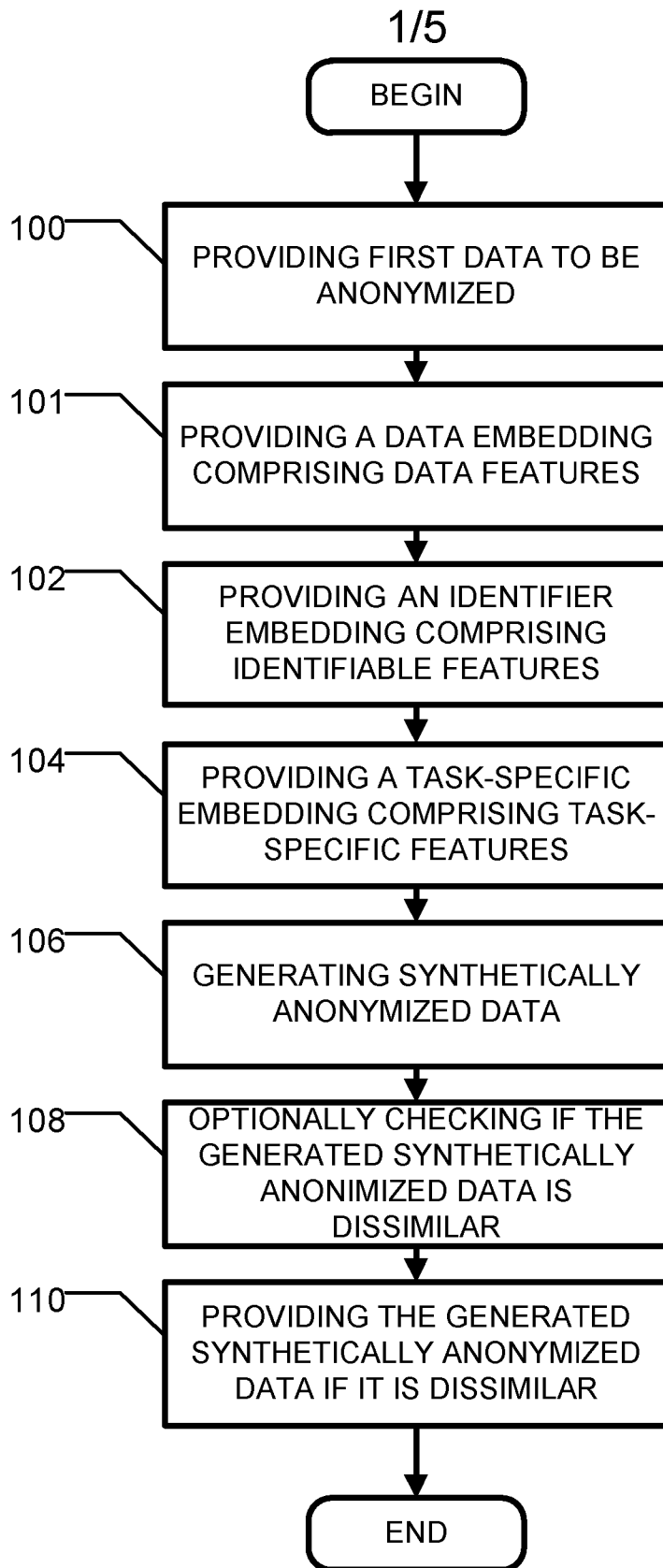


FIG. 1

2/5

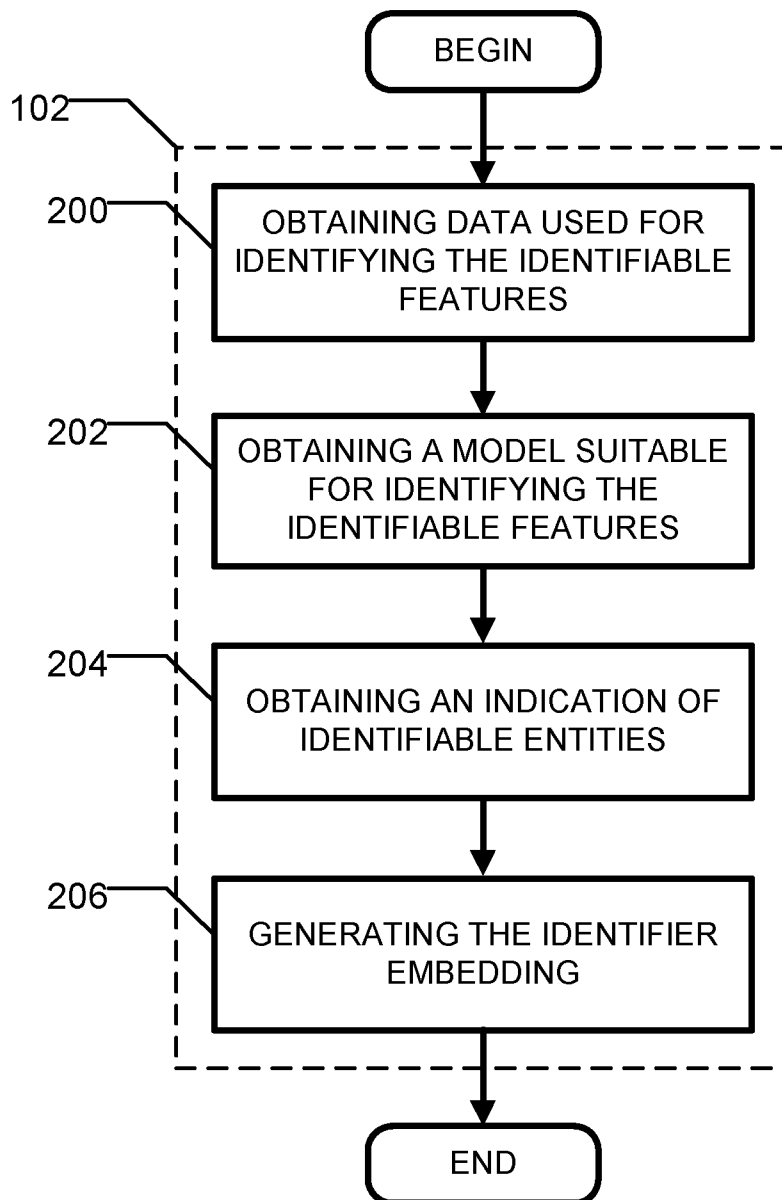


FIG. 2

3/5

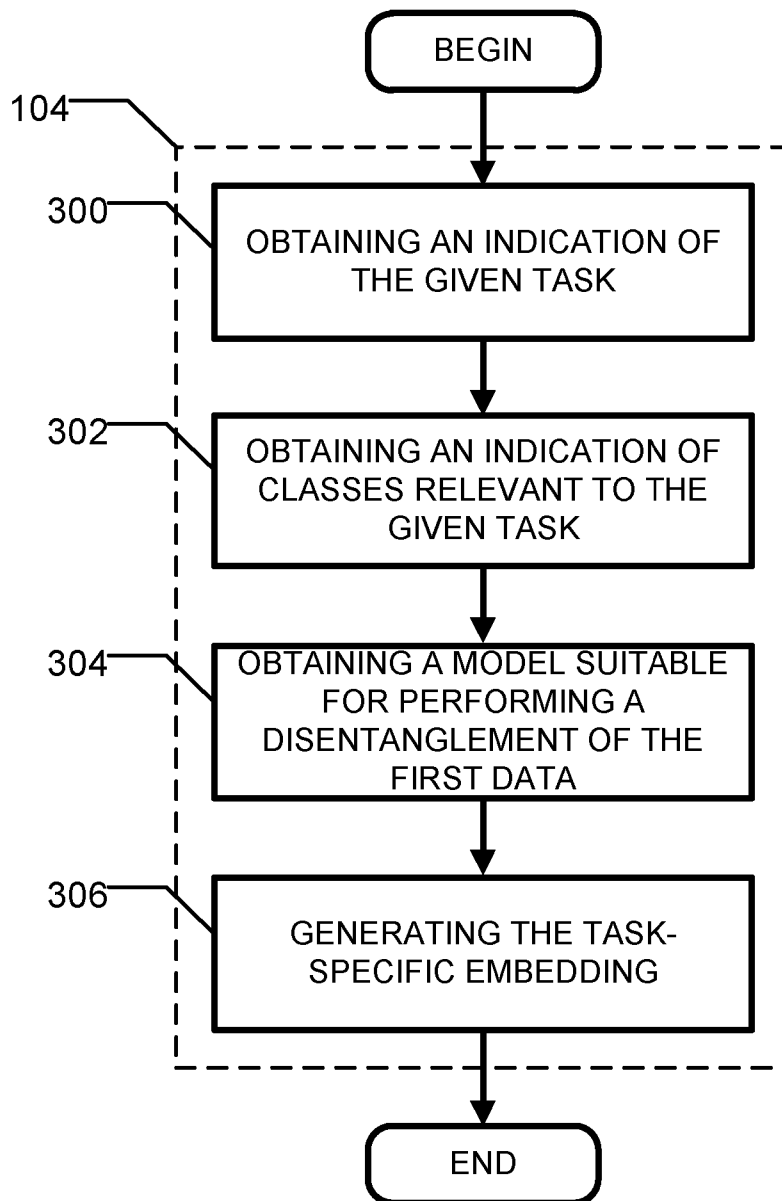


FIG. 3

4/5

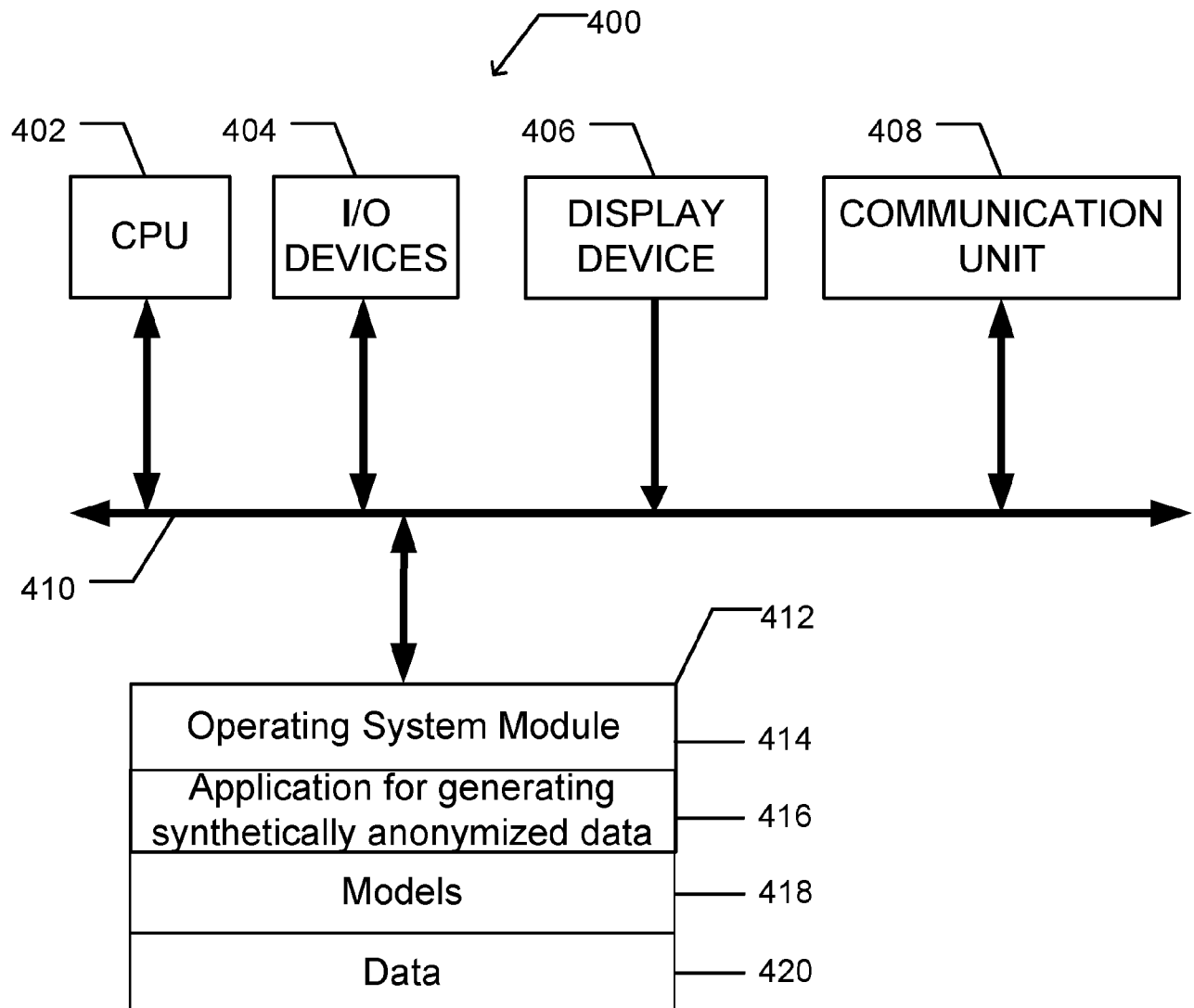


FIG. 4

5/5

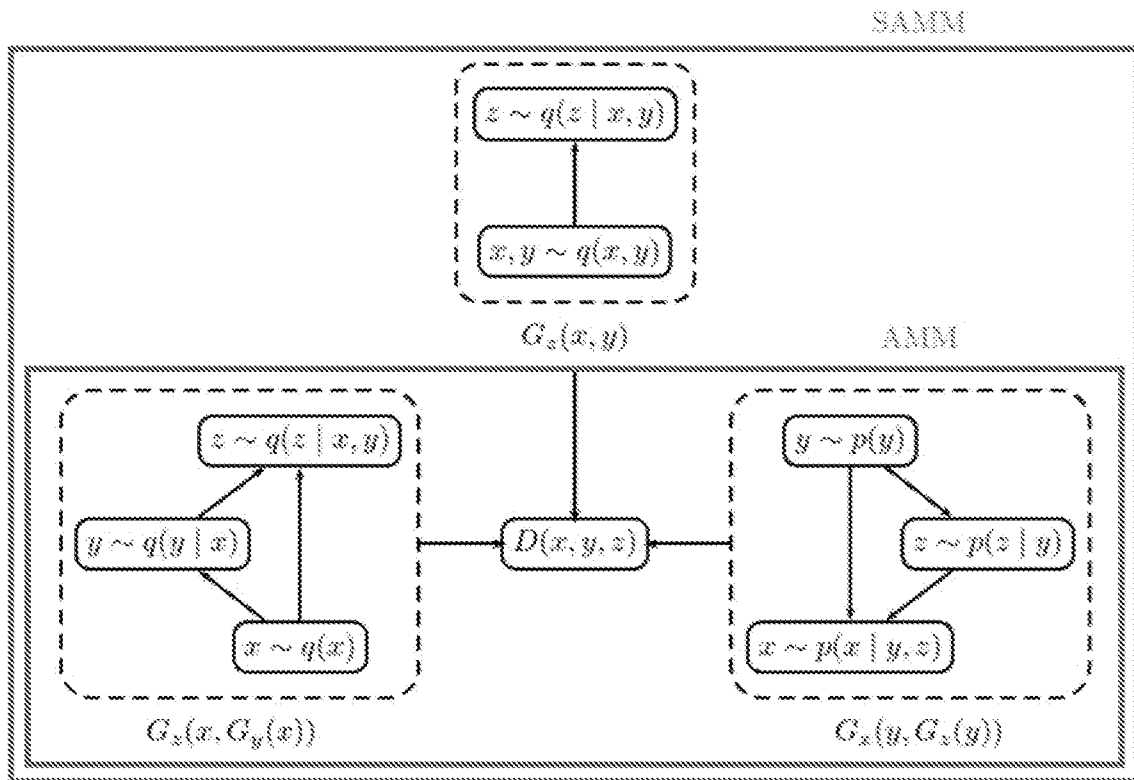


FIG. 5

BEGIN

100 PROVIDING FIRST DATA TO BE  
ANONYMIZED

101 PROVIDING A DATA EMBEDDING  
COMPRISING DATA FEATURES

102 PROVIDING AN IDENTIFIER  
EMBEDDING COMPRISING  
IDENTIFIABLE FEATURES

104 PROVIDING A TASK-SPECIFIC  
EMBEDDING COMPRISING TASK-  
SPECIFIC FEATURES

106 GENERATING SYNTHETICALLY  
ANONYMIZED DATA

108 OPTIONALLY CHECKING IF THE  
GENERATED SYNTHETICALLY  
ANONYMIZED DATA IS  
DISSIMILAR

110 PROVIDING THE GENERATED  
SYNTHETICALLY ANONYMIZED  
DATA IF IT IS DISSIMILAR

END