



(51) International Patent Classification:
G06F 21/62 (2013.01)

(21) International Application Number:
PCT/EP2021/070942

(22) International Filing Date:
27 July 2021 (27.07.2021)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **NEC LABORATORIES EUROPE GMBH**
[DE/DE]; Kurfürsten-Anlage 36, 69115 Heidelberg (DE).

(72) Inventors: **CIRILLO, Flavio**; NEC Laboratories Europe GmbH, Kurfürsten-Anlage 36, 69115 Heidelberg (IT).
CHENG, Bin; NEC Laboratories Europe GmbH, Kurfürsten-Anlage 36, 69115 Heidelberg (DE).

(74) Agent: **ULLRICH & NAUMANN**; Schneidmühlstraße 21, 69115 Heidelberg (DE).

(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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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,

(54) Title: METHOD AND SYSTEM FOR ENFORCING SECONDARY USAGE CONTROL ON DATA ANALYTICS SERVICE

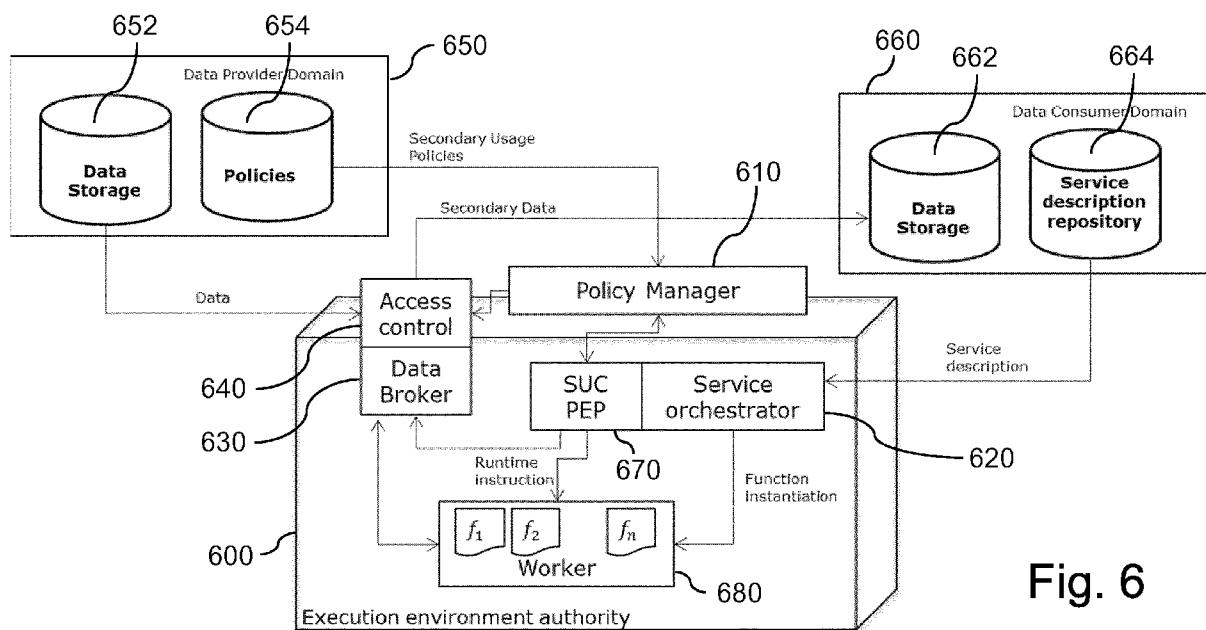


Fig. 6

(57) Abstract: Disclosed is a system and computer-implemented method of enforcing secondary data usage control. In order to enforce secondary usage control on data in a proactive and preventive manner without any direct interaction between data consumers and data providers, the method comprises providing, via a policy manager (610), secondary data usage policies of a data owner (650) of original data; providing, via a service orchestrator (620), a description of a service intended to be applied by a data consumer (660) to input data contained in the original data, the service including one or more data processing functions; matching, by a secondary usage control policy enforcement point, SUC PEP, component (670) the secondary data usage policies provided via the policy manager (610) with the input data or specific classes of the input data and/or with data processing functions of the service provided via the service orchestrator (620); and applying, by the SUC PEP component (670), the matched secondary usage policies on the secondary data.

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

Published:

— *with international search report (Art. 21(3))*

METHOD AND SYSTEM FOR ENFORCING SECONDARY USAGE CONTROL ON DATA ANALYTICS SERVICE

5 The present invention generally relates to the technical field of data protection. More specifically, the invention relates to a computer-implemented method and a system for enforcing secondary data usage control.

10 Data sharing between companies is fostered by both data consumers and data providers for mutual benefits to reduce costs for new services development and create new data business models.

15 Sharing infrastructures and platforms that are already available in the market are not yet generating an automatic data marketplace on a global scale. This is due to the data providers' reluctance to lose control over their owned data after data is shared. Methods for data usage control have been proposed to solve the *direct usage of data* by a data consumer (as shown in Fig. 1 between data provider/owner 100 and a first data consumer 110, i.e. company A), but *secondary usage control* (as shown in Fig. 1 between the first data consumer 110 and a second data consumer 120, i.e. company B) is still an open challenge regarding the control by the original data provider 100 over the secondary data processed by the first data consumer 110. The processed data can still carry sensitive information of the original data provider 100 or it can still be economically linked to data provider business. The General Regulation of Data Protection (GDPR, for reference, see, e.g., Voigt, P., von dem Bussche, A., 2017. The EU General Data Protection Regulation (GDPR). A Practical Guide. Springer Int. Publ. doi:10.1007/978-3-319-57959-7) also covers this aspect and hinders the business development of companies that consume personal data due to the uncertainty of the consequences if their processed data is shared or sold to third parties.

30 It is therefore an object of the present invention to improve and further develop a method and a system of the initially described type in such a way that secondary usage control on data is enforced in a proactive and preventive manner without any direct interaction between data consumers and data providers.

In accordance with the invention, the aforementioned object is accomplished by a computer-implemented method of enforcing secondary data usage control, the method comprising providing, via a policy manager, secondary data usage policies of a data owner of original data; providing, via a service orchestrator, a description of a service intended to be applied by a data consumer to input data contained in the original data, the service including one or more data processing functions; matching, by a secondary usage control policy enforcement point, SUC PEP, component the secondary data usage policies provided via the policy manager with the input data or specific classes of the input data and/or with data processing functions of the service provided via the service orchestrator; and applying, by the SUC PEP component, the matched secondary usage policies on the secondary data.

Furthermore, the aforementioned object is accomplished by a system for enforcing secondary data usage control, the system comprising a policy manager configured to provide secondary data usage policies of a data owner of original data; a service orchestrator configured to provide a description of a service intended to be applied by a data consumer to input data contained in the original data, the service including one or more data processing functions; and a secondary usage control policy enforcement point, SUC PEP, component configured to match the secondary data usage policies provided via the policy manager with the input data or specific classes of the input data and/or with data processing functions of the service provided via the service orchestrator, and apply the matched secondary usage policies on the secondary data.

The present invention focuses on the secondary usage control, that is the control on how the processed data produced by a data consumer can be used by a third party complying with the original data provider's policies. Secondary data usage control can happen through legal means rather than technical means as specified in the context of the present invention. In the case of legal enforcement, before any data analytics application is executed, the data consumer negotiates with the data owner for the usage of the data. The data owner, then, expresses his consent on the usage of the data under constraints. The constraints are usually encoded into contracts or policies. In contrast to such legal approach, the present invention

proposes a system and a method to enforce in a proactive and preventive manner secondary usage control on input data (specified by a data consumer and contained in the original data) without the direct interaction between data consumers and data providers or, more generally, without the direct intervention of humans in the loop.

5 The advantage is a scalable ecosystem for any kind of data sharing platform with a high degree of automation.

The present invention enforces the control of processed data according to the policies specified by the original data owner. The enforcement happens on services

10 without the involvement of human negotiations and in a preventive and proactive manner (no data is processed if not specified into the policies). Embodiments of the invention enforce the secondary data usage control by a centralized authority or by a federation of authorities (herein sometimes denoted execution environment authority). Embodiments of the invention assume that these authorities are trusted

15 environments that are certified and cannot be tampered (e.g., remotely attested).

According to some embodiments, the present invention provides methods and systems for applying secondary data usage policy that comprise a policy manager holding secondary data usage policies representing the will of the data owners. The

20 solution may further include a storage of data consuming applications in the form of composition of analytics tasks by data consumer. The analytics tasks may be defined by the input data or one or more particular classes of the input data, processing function, and output data.

25 According to an embodiment, the methods and systems further include matching secondary usage policy with the input data or particular classes of the input data and/or with processing functions of the service. The matched secondary usage policy may be applied on the output data by first deciding and then executing atomic actions that enforces:

- 30
- a. creating new access or data usage control policies targeting the output data;
 - b. instructing the system executing the analytics with commands; and/or
 - c. instructing the component handling the output data.

In the context of above item a.), the present invention provides a method for a data provider to specify policies on original data to be used as policies generator patterns to control the access and usage of any secondary (processed) data derived from the data provider's original data.

5

According to an embodiment, the present invention provides a method that enforces access and usage policies on secondary (processed) data in a proactive and preventive manner by deciding on and executing atomic actions on the policy manager (e.g., creation of new access, direct usage, and secondary policies), the data manager (e.g., deleting data), and the service orchestrator (e.g., instantiate pre-processing functions). The atomic actions may target the output data generated by services whose description (e.g., input data or class of input data and functions) matches with attributes of the policies.

10

15

In the context of the present disclosure, a class of input data may denote a specifically defined subset of the input data. For instance, the input data may be patient data including, e.g., name, address, age, gender, weight and disease(s) of the patients. In this case, each of the listed items may constitute a specific class of input data. In this example, privacy preserving policies may be defined that grant a data consumer, e.g., access to all classes except for the classes 'name' and 'address'.

20

According to an embodiment, an atomic action may include generating a policy for each set of output data specifying that the respective set cannot be accessed by any third parties except the original data owner and the data consumer. The generated policy may be stored into the policy manager storing.

25

According to a further embodiment, an atomic action may include generating a command routine for at least one processing node that executes the processing functions of the service to alter the execution of the service. The command routine may be send to an execution environment of the service. For instance, the command routine may be configured to stop and restart a container instance of the service applied to the original data at predefined time intervals.

30

According to a further embodiment, an atomic action may include modifying the service by prepending a processing function to the original data or appending a processing function to the processed (i.e. secondary) data. For instance, the service may be modified by anonymizing the input data by means of an anonymization function and applying the processing functions of the service to the anonymized data.

According to an embodiment, each of the data analytics functions is described by the data consumer with the input data or one or more particular classes of the input data, the processing function, and output data.

According to an embodiment, each of secondary data usage policies defines the owner of the original data, the input data or one or more particular classes of the input data targeted by the policy, the data consumer to whom the policy applies, at least one specific function targeted by the policy and at least one constraint specifying limitations on the usage of the targeted data.

According to an embodiment, the system is matching a policy with a data analytics function description by:

- i) matching the original data specified in the policy with the input data of class(es) of input data of the analytics function description;
- ii) matching the analytics function targeted by the policy with the data analytics function used by the service; and
- iii) matching the data consumer identity targeted by the policy with the data consumer identity that instantiates the data analytics service.

There are several ways how to design and further develop the teaching of the present invention in an advantageous way. To this end it is to be referred to the dependent claims on the one hand and to the following explanation of preferred embodiments of the invention by way of example, illustrated by the figure on the other hand. In connection with the explanation of the preferred embodiments of the invention by the aid of the figure, generally preferred embodiments and further developments of the teaching will be explained. In the drawing

Fig. 1 is a schematic view illustrating direct usage and secondary usage control of data of a data provider,

Fig. 2 is a schematic view illustrating first data usage control enforcement according to prior art,

Fig. 3 is a schematic view illustrating secondary data usage control enforcement according to prior art,

Fig. 4 is a schematic view illustrating a secondary data usage control system in accordance with an embodiment of the present invention,

Fig. 5 is a schematic view illustrating a secondary data usage control policy enforcement system in accordance with an embodiment of the present invention,

Fig. 6 is a schematic view illustrating a secondary data usage control system with a centralized authority in accordance with an embodiment of the present invention,

Fig. 7 is a schematic view illustrating a secondary data usage control system with a federation of authorities in accordance with an embodiment of the present invention,

Fig. 8 is a schematic view illustrating secondary data usage control enforcement on digital twins in accordance with an embodiment of the present invention,

Normal data usage control policy affects the access control and the way it is used by a first consumer analytics, as schematically shown in Fig. 2. Secondary usage control, instead, affects the enforcements on the resulting data of the first analytics, as schematically shown in Fig. 3.

A solution to secondary data usage control usually goes through private negotiations between data owner and data consumers with signed agreements followed by ad-hoc data sharing infrastructure deployment. Each data consuming service is then configured by an administrator to comply with the usage agreements. Controlling the actual enforcing of the agreed policies is done through audits that, however, would spot infringements only after they happened.

To address these issues, embodiments of the present invention provide systems and methods that enforce in a proactive and preventive manner secondary usage control on data without the direct intervention of humans in the loop. The advantage is a scalable ecosystem for a data sharing platform.

Fig. 4 illustrates the basic workflow of a method to protect *secondary data* (i.e. processed data from original data) in accordance with an embodiment of the present invention.

According to one aspect (denoted with an encircled 1 in Fig. 4), a data provider 400 holding its original data 402 submits secondary data usage policies protecting its owned data 402 to a policy manager 410. A secondary usage policy might refer to secondary data access, secondary data usage, and/or processed data from secondary data (i.e. transitivity).

According to a further aspect (denoted with an encircled 2 in Fig. 4), a data consumer 420 submits a description of the intended service that will process the original data to a service registry 430. The description of the intended service may be provided in the form of a composition of atomic analytics tasks 422. Each analytics task may be defined by the input data or a specific class of input data, a processing function, and the resulting output data.

According to a further aspect (denoted with an encircled 3 in Fig. 4), a policy enforcement logic 440 decides on atomic actions to be executed. In this context, it may be provided that the policy enforcement logic 440 is configured to match the secondary usage policies (as submitted by the original data owner 400 to the policy manager 410) with the input data or a specific class of input data and/or with

processing functions of the service (as submitted by the data consumer 420 to the service registry 430).

5 Based thereupon, the policy enforcement logic 440 may be configured to apply the matched secondary usage policies on the service output data by deciding on and executing atomic actions, such as creating access or data usage control policies targeting the output data (e.g., a policy stating that service output data cannot be shared with third parties), commanding a certain behaviour to computing nodes executing the analytics (e.g., flush service memory every two hours, instantiate a pre-processing anonymization task). Alternatively, the policy enforcement logic 440 may be configured to apply the matched secondary usage policies on the output data by commanding a certain behaviour to a component managing the service output data (e.g., service output data must be forwarded to original data owner as soon it is generated).

15 According to an embodiment of the present invention, the data provider 400 may specify secondary data usage policies in the following form that is stored into the policy manager 410:

- 20 - Who: I.e., indicating the owner of the data.
- What: I.e., indicating the data targeted by the policy, e.g., the type of data.
- To whom: I.e., indicating the data consumer that targets the policy. This information can be either a specific data consumer identifier or *any*.
- Purpose: I.e., indicating a specific function targeted by the policy. The function can be either a data consuming service name (denoted S hereinafter), a specific atomic task function (named $f(\cdot)$ hereinafter), or *any*.
25 - Constraints: I.e., indicating the limitations on the usage of the targeted data. The limitations may be formed by one or more rules. In this context, a rule is a combination of an *action type* to be enforced that is configured by *action parameters*.

30 According to an embodiment of the invention, a constraint on secondary data may specify a number of different control aspects, including access, data usage, and/or secondary usage (i.e. transitivity). With respect to the access control, it may be provided that the data owner defines secondary data access before the secondary

(i.e., processed) data is generated. For instance, access control may specify that processed data can be accessed (e.g., read) only by the data consumers that generated such processed data. The data usage control may specify how the secondary data can be used by a third party, for instance that processed data can be used only if aggregated with other datasets. The secondary usage control might affect all the processed data that directly or indirectly comes from the original data. For instance, secondary usage control may define that all processed data that descend from the original data must be accessible (access control) by the original data owner.

Policies on the secondary data may include, but are not limited to the following examples:

- Secondary data may be read by third parties, e.g. a third party may visualize the secondary (processed) data on a dashboard.
- Secondary data may be further processed by third parties, e.g. secondary processed data may flow into any analytics service of a third party.
- Secondary data may be processed only for a specific purpose, e.g. secondary data can be processed only by a specific service (independently from the third party requesting it).
- Secondary data must go through pre-processing, e.g. before flowing into any analytics service the secondary data must be aggregated with other datasets originated from other providers.
- Secondary data must go through post-processing, e.g. the final output of an analytics service using the secondary data (for example a time series) must be processed by an aggregation function (for example accumulate data in time windows of 10 minutes and only retrieve an average of the accumulated data).
- Secondary data can be used only within the computing premises of the original data provider, e.g. in order for an analytics service to use secondary data it must be executed on a processing nodes directly administrated by the original data provider.
- Secondary data can always be accessed by the original data provider, e.g. any processed data by third parties derived from the original data may flow into any analytics of the original data provider.

It is important to note that the original data provider/owner is not required to know the type of secondary data. For example, a data owner sharing the position of her smartphone might specify that any processed data that uses her position information cannot be used by any marketing purposes. As a more concrete example, a data owner shares her position to a navigation system that will use it to monitor the road traffic to better estimate a suggested route. However, the data owner may specify that any traffic information using her data cannot be used by a marketing analytics (e.g. with the purpose of optimal advertisement sign planning on the roadside). The original data owner does not need to specify the traffic information type into her policy since data usage policies will be automatically generated for each service description that will use her data.

According to embodiments of the invention, with regard to the service specification by a data consumer it may be provided that a service is designed as a combination of processing functions. The services may be stored into a service description repository (e.g. in service description repositories 664 or 764, as shown in Figs. 6 and 7, respectively). A processing function $f(\cdot)$ of a service may be specified as:

$$(y_1, y_2, \dots, y_m) = f(x_1, x_2, \dots, x_n),$$

where x_i denotes some input data (contained in the original data) and y_i denotes outputs of secondary/processed data.

Instead of using original data as inputs for processing function $f(\cdot)$, the output of another function can be used. For example:

$$(y_1, y_2, \dots, y_m) = f(x_1, g(\cdot), \dots, x_n)$$

The original input set of function $g(\cdot)$ and of function $f(\cdot)$ can overlap. Function $g(\cdot)$ might be also a combination of other function.

As already mentioned above, according to embodiments of the invention, a service S can be designed as a combination of functions:

$$S: Y_s = f_s(\cdot)$$

5 The input x_i univocally identifies an original data element of a data provider such as a specific data set or a particular class of data (e.g., surveillance video data) or a specific instance of a class of data, or a processed data element from a function instance. The output y_j univocally identifies a processed data element from the particular function instance. In some embodiments, the unique identifiers are handled automatically by the system.

10 According to embodiments of the invention, the enforcement of the secondary usage policies (specified by a data owner) on a service S (specified by a data consumer) is done by matching the secondary usage policies with input data and/or with processing functions of the service S . Specifically, matching the policies may include at least one of the following steps or any combinations thereof:

- Matching the “what” of the policies with x_i , that is the input data or the class(es) of input data of all the functions of the combination of the service S .
- Matching the “purpose” of the policies with the service S name or its composing functions $f_s(\cdot)$
- Matching the “to whom” of the policies with the data consumer identifier requesting the service S .

25 According to embodiments of the invention, enforcing secondary usage policies on the service S may further include applying the constraints specified by the policies to the output Y_s of the service S . This may be done by deciding on and executing atomic actions. For instance, in an embodiment the method provides for generating one policy for each $(y_1, y_2, \dots, y_m)$. Each policy may state that y_i cannot be accessed by any third parties (i.e., all third parties except original data owner and data consumer). For execution, the generated policy may be stored into a policy manager (e.g., policy manager 410 shown in Fig. 4). In another embodiment, a commands routine may be generated and send to an execution environment of the service S that stops and restarts a container instance of the service S in predefined time intervals, e.g. every two hours. In yet another embodiment, an anonymization

function $f_a(\cdot)$ may be prepended to the service S , i.e. the service S is modified by adding $f_a(\cdot)$ to the function using the original data. For instance, if x_1 is the original data input to be anonymized, $S : (y_1, y_2, \dots, y_m) = f(x_1, x_2, \dots, x_n)$ is changed to $S : (y_1, y_2, \dots, y_m) = f(f_a(\cdot x_1), x_2, \dots, x_n)$. In yet another embodiment, a subscription may be generated that targets Y_s with notification callback set as the component managing data of the original data provider. For executing this policy, the generated subscription may be instantiated to the component managing Y_s . The above mentioned atomic actions may be decided and executed in isolation or in any combination with each other.

In some embodiments, a constraint might enforce the creation of an access control policy on the output y_s (e.g., not share with other parties, access granted to original data owner, etc.). In this case, as shown in Fig. 5, a Secondary Usage Control policy enforcement point (SUC PEP) component 500, which is configured to receive policies from a policy manager 510 and data analytics service description from a service description repository 520, may generate one atomic action on the output y_s . This atomic action may be executed by storing the newly generated policy. This method implements an *automatic policies generator*.

In other embodiments, the constraint might enforce the creation of data usage control policy on the output y_s (e.g., aggregate y_s with other data before further use). Also in this case, the SUC PEP 500 may decide to generate new policies covering the data usage control covering y_s that will be stored.

Some embodiments might instruct the execution environment to apply runtime commands/instructions (e.g., delete output data y_s in regular time intervals, for instance every two hours). In this case, the SUC PEP 500 may generate a routine with a timer into the execution environment to flush the process memory of the analytics service (e.g., by stopping and restarting a container).

Application scenario 1: Centralized authority

Fig. 6 illustrates an embodiment of the present invention where the secondary data usage control is enforced by a centralized authority 600. The centralized authority 600 holds the data and the execution environment where data is processed.

5

According to an embodiment, the centralized authority 600 may offer three interfaces: 1) A policy manager 610, where a data provider 650 can set policies 654, in particular secondary usage policies. 2) A service orchestrator 620 where a data consumer 660 can send the service description (e.g. from a service description repository 664). 3) A data broker 630 protected by an access control system 640 where a data provider 650 can submit its data (e.g. from data storage 654) and where data consumers 660 retrieve secondary data (e.g. to be stored in their data storage 662).

10

15

According to the illustrated embodiment, the centralized authority 600 further comprises a Secondary Data Usage Control Policy Enforcement Point (SUC PEP) 670 that may be triggered once a new service description is submitted to the service orchestrator 620. The SUC PEP 670 may be configured to use the secondary data usage control policies from the policy manager 610 to makes decisions. The decisions may be either to generate new policies (access, data usage or secondary usage control), or to alter the functioning of the data management system (i.e., in particular the Data Broker 630 shown in Fig. 6), to alter the service description by imposing data pre- or post-processing functions, or to alter the processing of the processing nodes, namely workers 680. After the decisions from the SUC PEP 670, the functions of the service, possibly modified by the SUC PEP 670, may be instantiated into the workers 680. To this end, the SUC PEP 670 may be configured to transmit respective run-time instructions to the workers 680.

20

25

30

The data consumer 660 might retrieve the secondary data if allowed by the data provider policies 654. Alternatively, the data consumer 660 or another consumer might use the secondary data for further processing, if the data provider policies 654 allow it.

Application scenario 2: Federation of authorities

Fig. 7 illustrates an embodiment of the present invention where the secondary data usage control is enforced by a federation of authorities 700. While the illustrated embodiment, for the sake of clarity, illustrates a federation of only two authorities, namely execution environment authority 700a located in the data provider domain and execution environment authority 700b located in the data consumer domain, it will be appreciated by those skilled in the art that in practical deployments the federation may include a higher number of authorities.

The exposed interfaces for the data consumers and data providers are the same as in the embodiment described above in connection with Fig. 6: i) a policy manager 710 where the data provider sets policies, ii) a service orchestrator 720a, 720b where the data consumer sends the service description, iii) a data broker 730a, 730b protected by an access control system 740a, 740b where a data provider submits its data and data consumers retrieve secondary data.

In the illustrated embodiment, the federated authorities 700a, 700b share the same policy manager 710. However, in some other embodiments, also this component can be federated. In yet some other embodiments, the policy manager 710 can be distributed.

According to the illustrated embodiment, a data provider sets policies 754 into the policy manager 710 and sends the original data to the data broker 730a of the local execution environment authority 700a. A data consumers submits the service description from its service description repository 764 to the service orchestrator 720b of its local execution environment authority 700b.

The two execution environment authorities 700a, 700b may be configured to exchange between each other the service descriptions in order to have a distributed decision on the policies among the SUC PEPs 770a, 770b. In some embodiments, the service descriptions might be stored into a centralized component such as the service registry similar to the policy manager 710. In such case, the data consumer

would have to submit only the service name to the local service orchestrator 720b to trigger the service. In other embodiments, this service registry might be a distributed component.

5

Application scenario 3: Digital Twin for Healthcare

Embodiments of the present invention fit very well for digital twin scenarios, for example a healthcare scenario, as schematically illustrated in Fig. 8. A digital twin is a digital representation of a real world object or person. Digital twins are linked with each other as objects and persons are linked in the real world. Digital twins are connected with each other through analytics services.

10

15

20

25

30

A prominent use case for e-health, which is schematically illustrated in Fig. 8, is the monitoring of disease spread and building management to minimize the risks of contagion within indoor environments. For example, traveling persons arriving at an airport might share their current location and travelling data to the airport infrastructure that monitors the contact tracing between persons. Airport administration computes the status of different areas of the airport and the possible encountering between persons. The airport is using the data from the travellers to estimate the usage of different types of common spaces (e.g., restauration areas with vending machines, resting areas, entertainment areas, restrooms). These estimations may be used by the building management to open or close common areas on the needs. For example, if at certain time of the day the number of people seeking for entertainment is decreasing, the total size of entertainment areas might be reduced, for example, by closing gates, reducing HVAC and switching off lights. Once closed, those areas can go through automatic sanitizing procedures to make the areas disinfected from disease agents. This optimal building management is advantageous for the airport since it minimizes the usage of energy for areas management, it minimizes the areas to be sanitized, and increase the control over the overall airport. On the other side, the travellers experience a better service since areas are opened on the needs and sanitized when less needed, and they are passing and stationing on areas with high standards of sanitization.

Travellers might agree on sharing their tracing information for the benefits they get back (e.g., better user experience), but usually they want to avoid that their information is used for marketing analysis. Thus, the travellers 800 can protect how a crowd estimation 820 derived from their data is used by registering a secondary (processing) data policy at a SUC PEP 870 stating, among other information, the purpose of data usage. Such policy might be easily configured by agreeing or disagreeing with a “terms and conditions” contract on their own smartphone when connecting to the local wireless network of the airport.

In other embodiments, the personal data might be covered by secondary usage control policies decided by the single persons or by regulation (e.g., GDPR). The secondary usage control enforcement depicted in Fig. 8 can be configured to affect the managing of secondary data (aka processed data), the linked digital twin, and third part services, as shown in Fig. 8.

Application scenario 4: Digital Twin for Industry 4.0

Data sharing technologies are an important enabler for industry 4.0 scenarios where producers and suppliers need to interact for efficient production.

Assuming the scenario depicted in Fig. 9, three industrial companies cooperate with each other. Company A has a chemical plant that produces basic chemicals for other processing companies, such as Company B. Company A and Company B are assumed to be located in the same industrial area since the transportation and storage of some materials are expensive, thus it is better to have them shipped and used as quick as possible to reduce costs.

Company B produces different kind of materials with its machineries. For a specific product B2, Company B uses machine B1 and awaits for the supply of a certain chemical stock A2 from Company A. Chemical stock A2 is assumed to be expensive to be stored for long time, thus it is better to have it used quickly.

As shown in Fig. 9, company A uses the data from the chemical machine A1 to predict when the chemical stock A2 is ready to be delivered. When the stock is ready it is transported through dedicated tubes or with autonomous driving trucks. The information of the predicted stock delivered is used by the processing machine
5 routine that commands processing machine B1 to automatically complete any production that machine B1 is doing in order to start the production of product B2.

The information of processing machine B1 is used by a centralized storage system to predict the needed supply by the processing plant of Company B to make the
10 supply order C1 to company C.

Company A is a specialized chemical company and it is assumed that they want to keep their production lines secret. Thus, by applying a method according to the present invention as disclosed herein, it can be enforced via the SUC PEP 970 that
15 the prediction of chemical stock A2 can be used by company B only to command their local machine (as indicated by the lower dotted line in Fig. 9).

Further, as indicated by the upper dotted line in Fig. 9, it can be enforced via the SUC PEP 970 that the supply prediction analytics can use the data from the
20 processing machine B1 only aggregated with data coming from other machineries. With the protection of this secondary data, company C cannot infer how much impact the supply from company A affects the supply order C1.

Without technical enforcement of the secondary data usage control disclosed by this
25 invention the information might be protected only by legal agreements that can be enforced only after a possible data leakage has been detected.

Many modifications and other embodiments of the invention set forth herein will come to mind to the one skilled in the art to which the invention pertains having the
30 benefit of the teachings presented in the foregoing description and the associated drawings. Therefore, it is to be understood that the invention is not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although

specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

C l a i m s

1. A computer-implemented method of enforcing secondary data usage control, the method comprising:
 - 5 providing, via a policy manager (610), secondary data usage policies of a data owner (650) of original data;
 - providing, via a service orchestrator (620), a description of a service intended to be applied by a data consumer (660) to input data contained in the original data, the service including one or more data processing functions;
 - 10 matching, by a secondary usage control policy enforcement point, SUC PEP, component (670) the secondary data usage policies provided via the policy manager (610) with the input data or specific classes of the input data and/or with data processing functions of the service provided via the service orchestrator (620); and
 - 15 applying, by the SUC PEP component (670), the matched secondary usage policies on the secondary data.
2. The method according to claim 1, wherein the description of the service is provided in the form of a composition of analytic tasks.
- 20 3. The method according to claim 2, wherein each of the analytics tasks is defined by the input data or specific classes of the input data, a data processing function, and output data.
- 25 4. The method according to any of claims 1 to 3, wherein secondary data usage policies define the owner (650) of the original data, the data targeted by the policy, the data consumer (660) to whom the policy applies, at least one function targeted by the policy and at least one constraint specifying limitations on the usage of the targeted data.
- 30 5. The method according to any of claims 1 to 4, wherein applying the matched secondary usage policies on the secondary data comprises deciding on and then executing atomic actions that enforce the creation of new access or data usage control policies targeting the output data.

6. The method according to any of claims 1 to 5, wherein applying the matched secondary usage policies on the secondary data comprises deciding on and then executing atomic actions that enforce instructions for a component (680) that executes the service applied to the original data and/or for a component that handles the secondary data.

7. The method according to any of claims 1 to 6, wherein an atomic action includes

generating a policy for each set of output data specifying that the respective set cannot be accessed by any third parties except the original data owner (650) and the data consumer (660); and

storing the generated policy into the policy manager.

8. The method according to any of claims 1 to 7, wherein an atomic action includes

generating a command routine for at least one processing node (680) that execute the data processing functions of the service to alter the execution of the service; and

sending the command routine to an execution environment of the service.

9. The method according to any of claims 1 to 8, wherein an atomic action includes

modifying the service by prepending a processing function to the original data or appending a processing function to the processed data.

10. A system for enforcing secondary data usage control, the system comprising: a policy manager (610) configured to provide secondary data usage policies of a data owner (650) of original data;

a service orchestrator (620) configured to provide a description of a service intended to be applied by a data consumer (660) to input data contained in the original data, the service including one or more data processing functions; and

a secondary usage control policy enforcement point, SUC PEP, component (670) configured to

match the secondary data usage policies provided via the policy manager (610) with the input data or specific classes of the input data and/or with data processing functions of the service provided via the service orchestrator (620), and apply the matched secondary usage policies on the secondary data.

5

11. The system according to claim 10, further comprising a data management system including a data broker (630) protected by an access control system (640) configured to receive the original data from the data owner (650) and to provide the secondary data to the data consumer (660).

10

12. The system according to claim 11, wherein the SUC PEP component (670) is configured to use the secondary data usage control policies from the policy manager (610) to makes decisions, wherein the decisions relate to at least one of generating new policies, altering the functioning of the data management system, altering the service description by imposing data pre- or post-processing functions, and altering the processing of the processing nodes (680) that execute the processing functions of the service.

15

13. The system according to claim 11 or 12, wherein the data management system, the policy manager (610), the service orchestrator (620) and the SUC PEP component (670) are part of a centralized execution environment authority (600).

20

14. The system according to 11 or 12, wherein the data management system, the service orchestrator (720a, 720b) and the SUC PEP component (770a, 770b) are constructed to build up a federation of execution environment authorities (700a, 700b).

25

15. The system according to claim 14, wherein the federated execution environment authorities (700a, 700b) share the same policy manager (710).

30

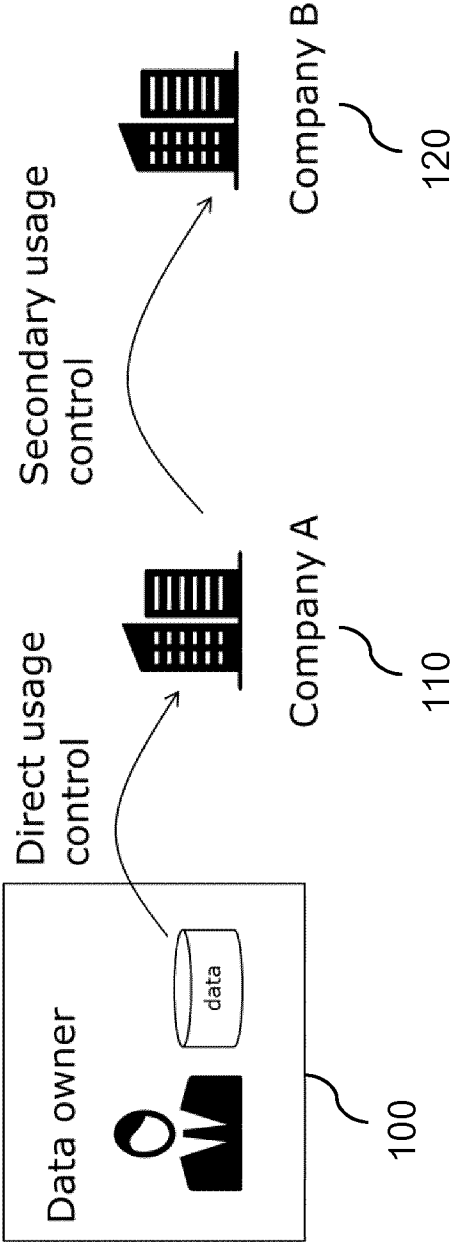


Fig. 1

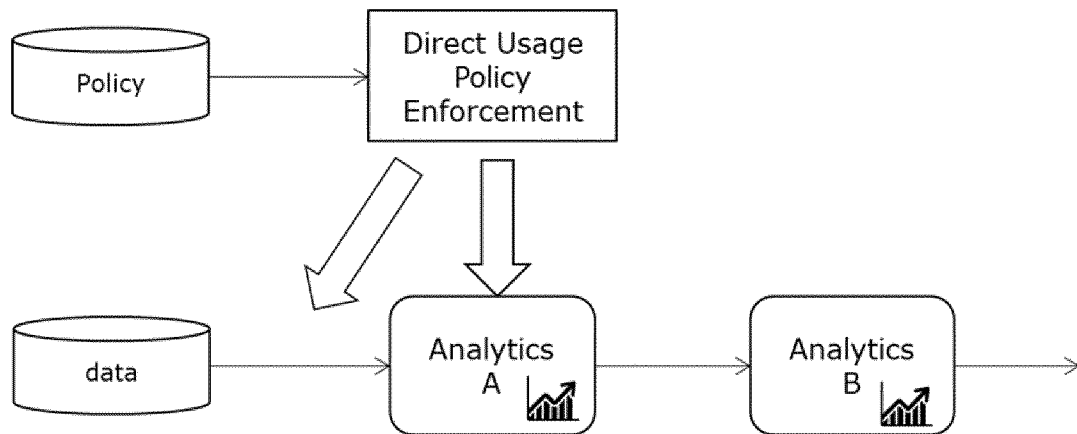


Fig. 2

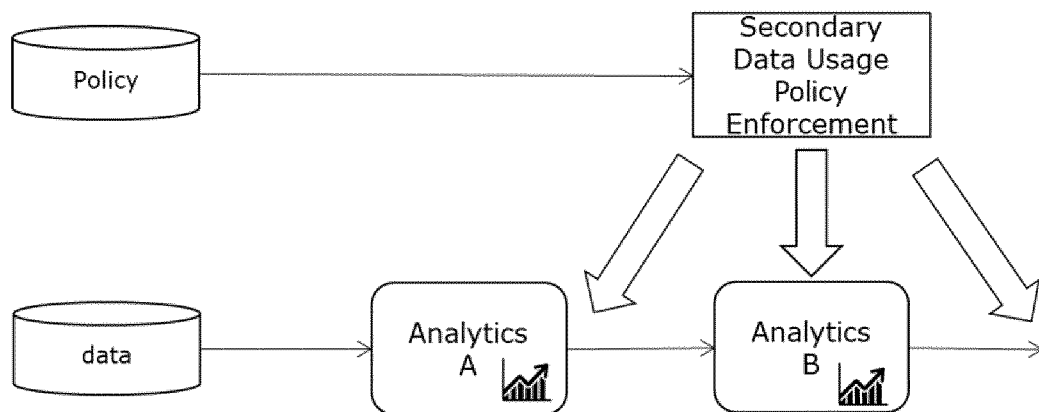


Fig. 3

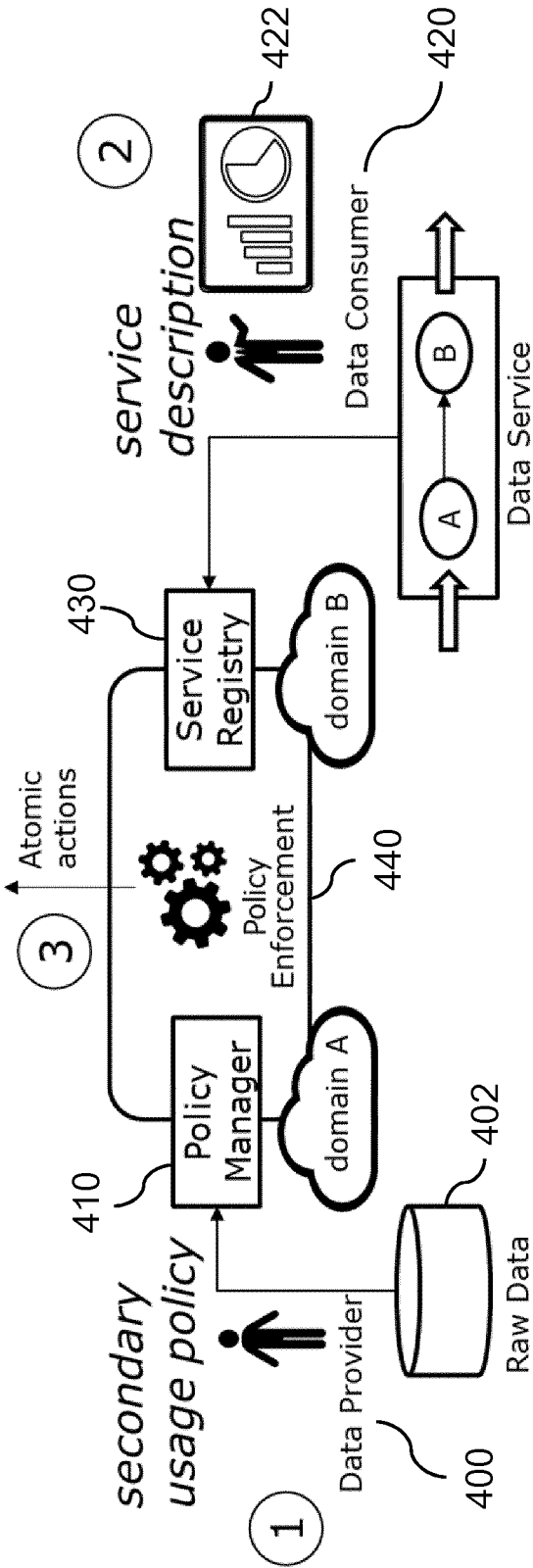


Fig. 4

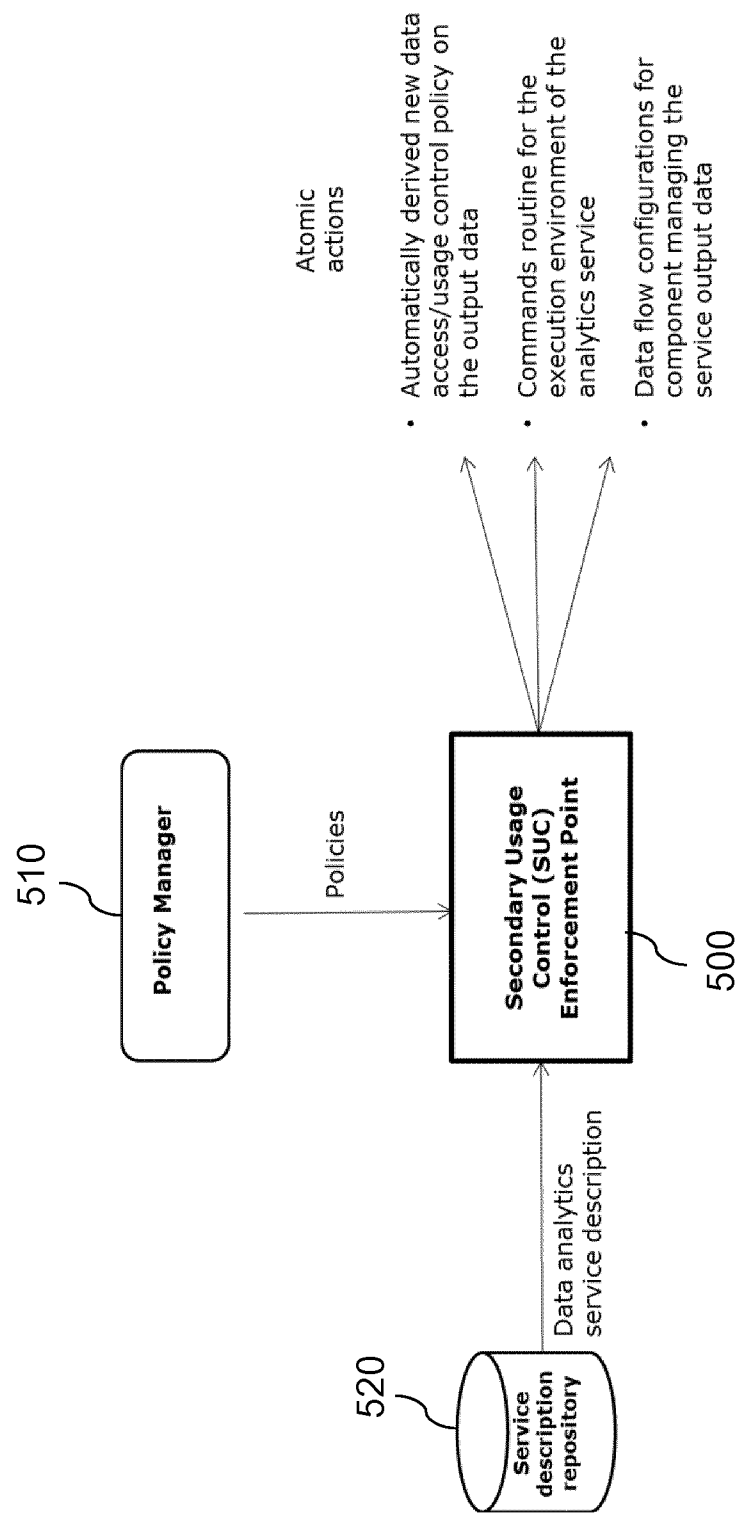


Fig. 5

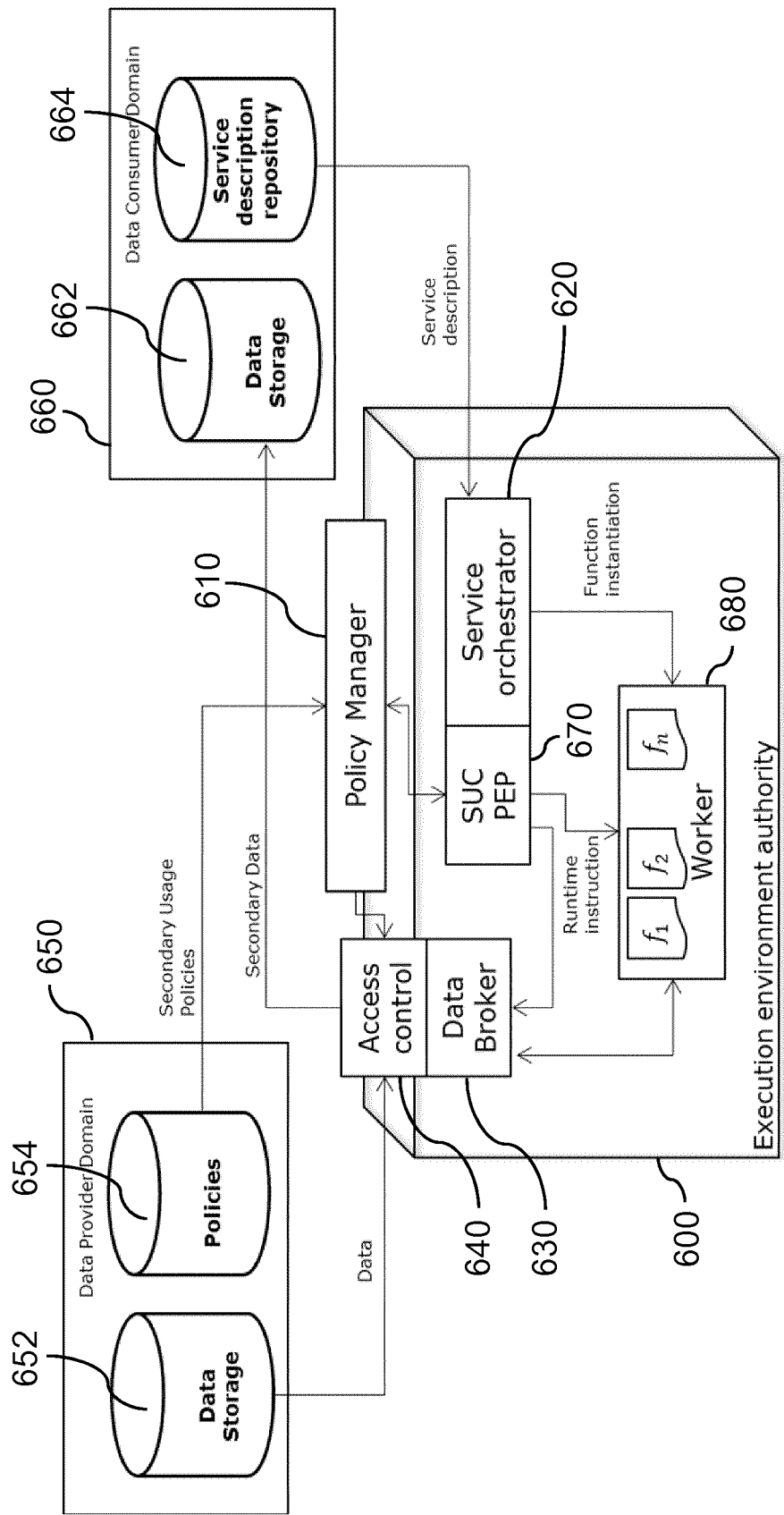


Fig. 6

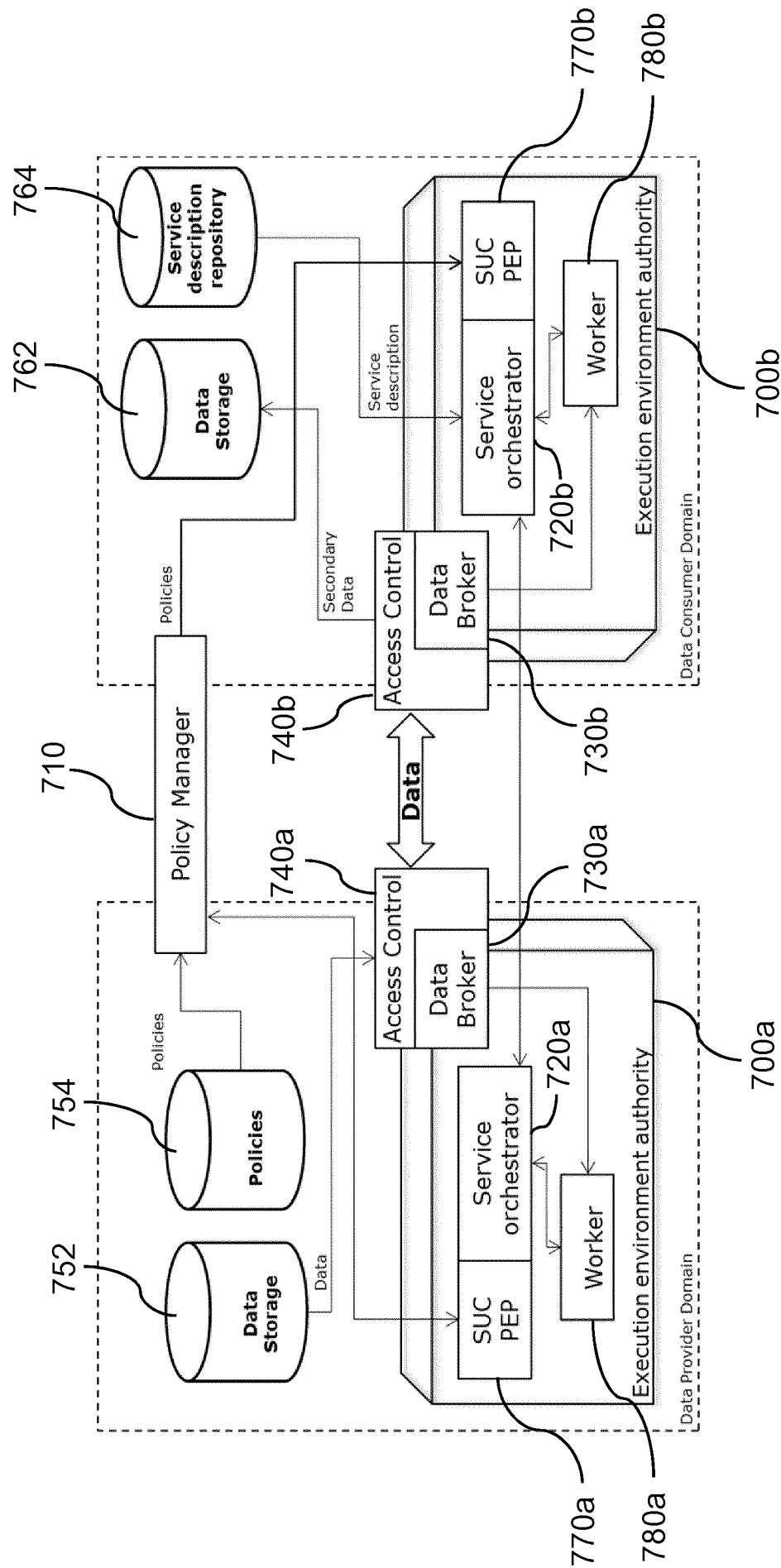


Fig. 7

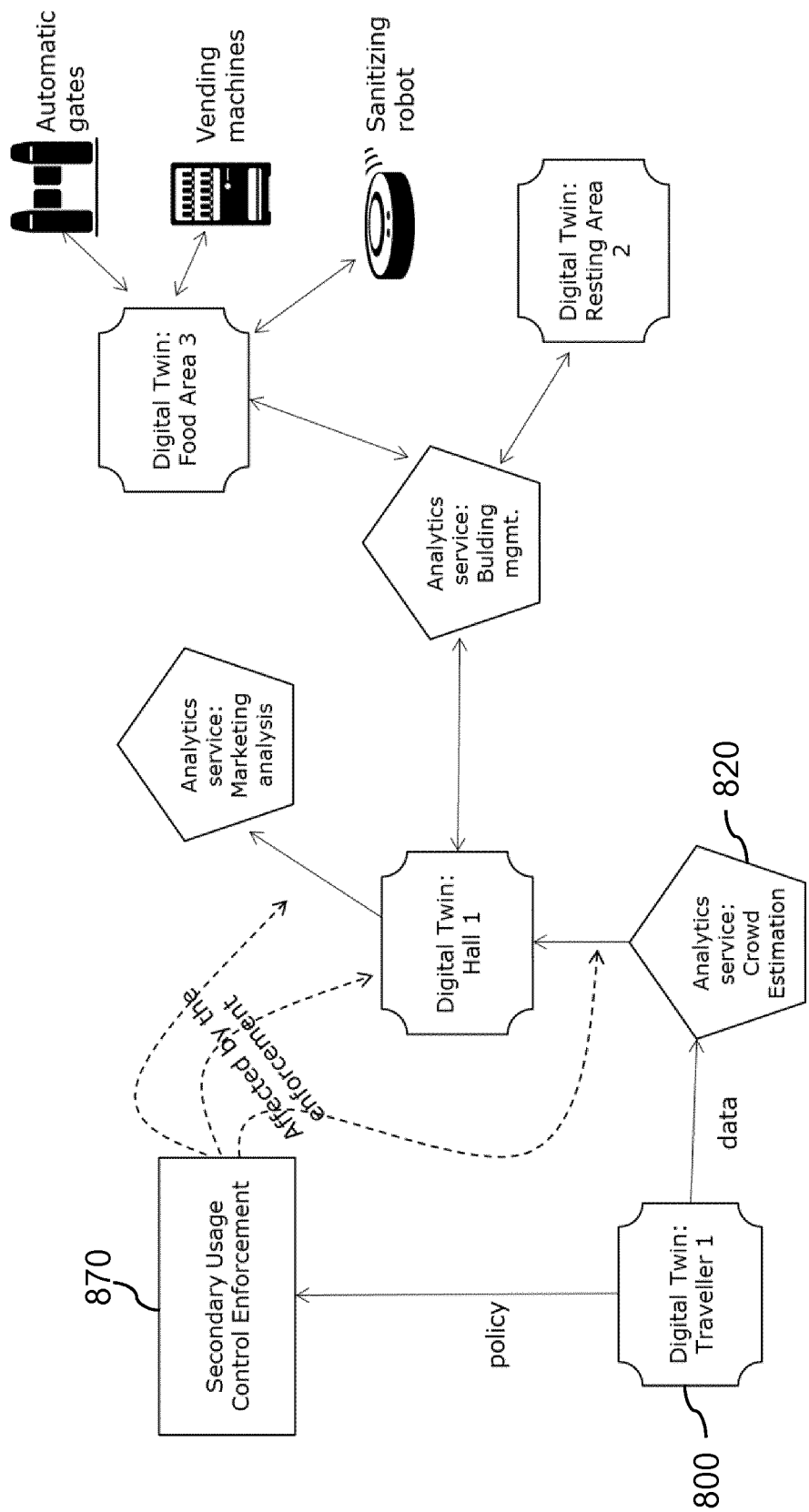


Fig. 8

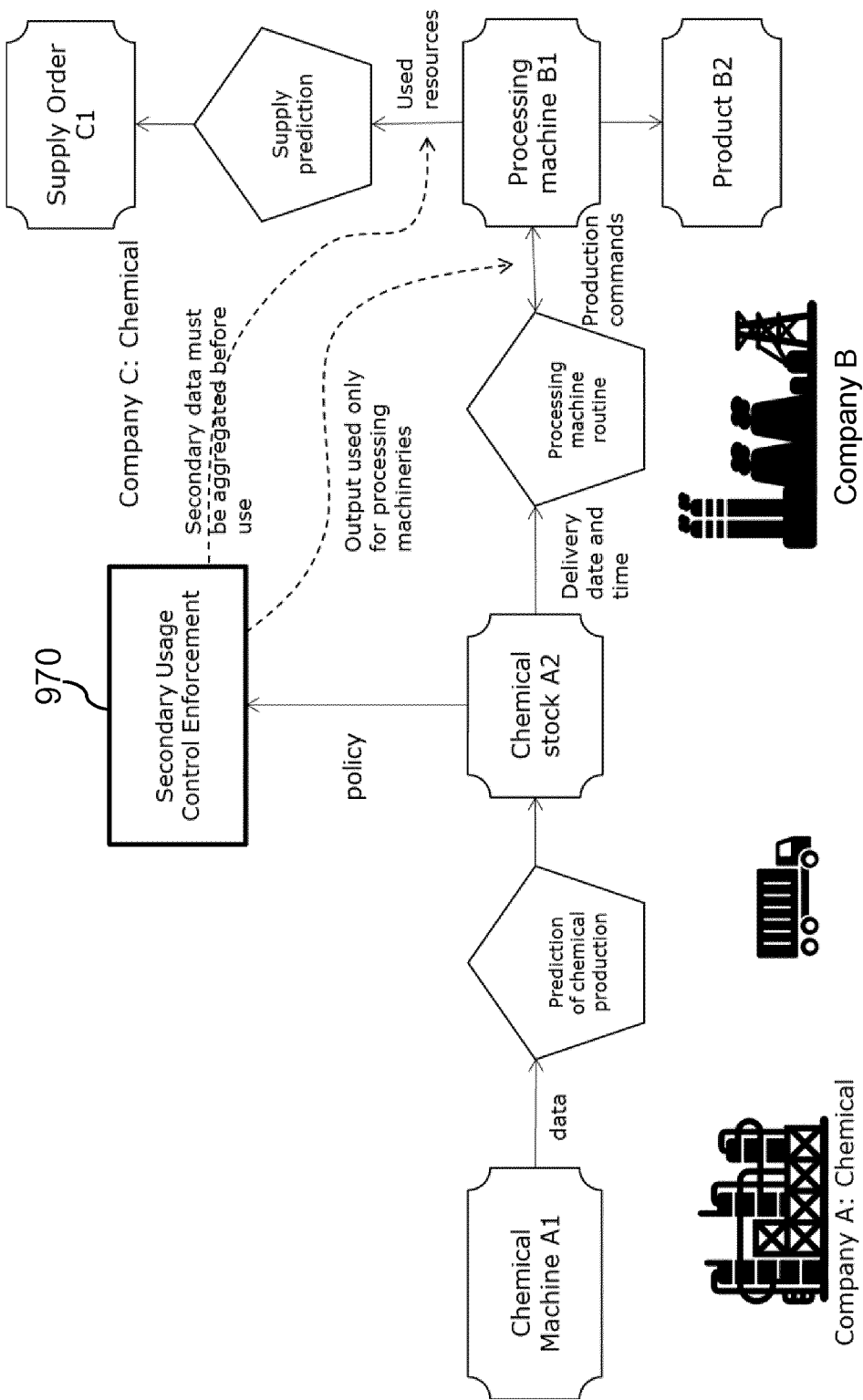


Fig. 9

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2021/070942

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F21/62 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F H04L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2010/205159 A1 (LI JUN [US] ET AL) 12 August 2010 (2010-08-12) paragraph [0001] paragraph [0009] - paragraph [0010] paragraph [0012] paragraph [0019] paragraph [0021] paragraph [0024] figures 4, 5 ----- -/--	1-15
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 16 March 2022		Date of mailing of the international search report 25/03/2022
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Bharucha, Zubin

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2021/070942

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020/004442 A1 (CASWELL PAUL CLIVE [GB] ET AL) 2 January 2020 (2020-01-02) paragraph [0001] paragraph [0010] paragraph [0012] paragraph [0015] paragraph [0018] - paragraph [0019] paragraph [0021] - paragraph [0024] paragraph [0032] paragraph [0040] - paragraph [0042] paragraph [0045] paragraph [0047] - paragraph [0048] figures 6A, 6B -----	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2021/070942

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010205159 A1	12-08-2010	NONE	
US 2020004442 A1	02-01-2020	CN 110692223 A	14-01-2020
		EP 3652653 A1	20-05-2020
		JP 6894985 B2	30-06-2021
		JP 2020521213 A	16-07-2020
		US 2020004442 A1	02-01-2020
		WO 2019013818 A1	17-01-2019