

DENGUE OUTBREAK PREDICTION

CLAIM OF PRIORITY

[0001] This application claims priority to U.S. Provisional Application No. 63/581,325, filed September 8, 2023, which is incorporated by reference in its entirety.

5 TECHNICAL FIELD

[0002] The present invention relates to a computer-implemented method of providing a dengue data platform for predicting an outbreak of dengue disease, also referred to herein as dengue outbreak. The present invention further relates to a corresponding computer program and to a computer-readable medium storing such a computer program. Moreover, the present invention relates to a
10 computing system for predicting a dengue outbreak.

BACKGROUND

[0003] Dengue disease or fever, commonly referred to herein as dengue, can be a major public health concern in many countries worldwide, particularly in tropical and subtropical regions, such as Latin America or Asia. The World Health Organisation, WHO, has been reporting an increasing
15 number of dengue cases over the past few decades, with an estimated minimum number of 50 million infections per year worldwide in recent years.

[0004] Due to increasing temperatures and changing weather conditions on a global scale, the number of dengue cases is expected to continue to rise in the coming years. Also, countries at lower geographical latitude, such as European or North American countries, are expected to be affected
20 by dengue disease in the near future.

[0005] Dengue disease is a vector-borne disease, which is caused by various dengue virus strains and spread by mosquitos. It is among the fastest-spreading virally-transmitted diseases, and therefore can lead to rapid outbreaks with rapidly increasing numbers of dengue cases in a short period of time, in particular in densely populated regions. Such outbreaks can impose a large burden
25 on public health and sometimes overwhelm local or national healthcare systems.

[0006] To protect public health, for example to prepare for potential dengue outbreaks and take corresponding countermeasures, numerous different dengue prediction or forecasting techniques have been developed and are currently used by numerous scientific working groups, agencies and institutions worldwide. However, differences in dengue prediction techniques and approaches can
30 lead to different predictions or estimates, wherein the differences in the quality and availability of the data utilised by the individual techniques may have the largest impact. For example, different techniques currently used in conjunction with data of differing quality and availability can lead to diverse predictions of both the severity and the time-evolution or timing of a dengue outbreak at a

given location. This can further lead to confusion and delays in taking appropriate countermeasures to minimize a potential dengue outbreak.

SUMMARY

[0007] It may, therefore, be desirable to provide for an improved dengue data platform, also referred to herein as “dengue platform” or similar, for predicting a dengue outbreak or dengue infection in general, which can overcome or at least mitigate the aforementioned drawbacks of current approaches or techniques used to predict dengue outbreaks.

[0008] This is achieved by the subject matter of the independent claims, wherein further embodiments are incorporated in the dependent claims and the following description.

[0009] Aspects of the present disclosure relate to (a) a computer-implemented method of providing a dengue platform for predicting one or more parameters associated with a dengue outbreak, (b) a computer program configured to instruct a computing system or dengue platform to perform steps of said method, (c) a computer-readable medium, for example a non-transitory computer-readable medium, storing such program, and (d) a computing system for predicting one or more parameters associated with a dengue outbreak. Any disclosure presented herein with reference to one or an aspect of the present disclosure equally applies to any other aspect of the present disclosure.

Aspects related to the computer-implemented method

[0010] According to an aspect, there is provided a computer-implemented method of providing a dengue platform for predicting one or more parameters and/or precursors associated with a dengue outbreak and/or for predicting a dengue outbreak. Alternatively or additionally, the method may relate to a computer-implemented method of operating a computing system and/or dengue platform for predicting one or more parameters and/or precursors associated with a dengue outbreak. It is noted that the method is not limited to the prediction of one or more parameters and/or precursors associated with a dengue outbreak, but can equally be used to predict or estimate one or more parameters and/or precursors related to dengue infection in general. The method comprises the following steps:

- collecting, in a raw data repository, raw dengue case data from a plurality of different dengue case data sources, the raw dengue case data being indicative of one or more dengue cases reported at one or more geographical locations for one or more time periods, which can include one or more time instants;
- collecting, in a raw data repository, raw weather data from one or more weather data sources, the raw weather data being indicative of one or more weather parameters at one or more geographical locations and one or more time periods or time instants;
- generating a refined data repository based on consolidating the raw dengue case data and the raw weather data from the raw data repository corresponding to one or both geographical location and time period;

- receiving, via an application programming access interface (“API”), an API call requesting a dengue modelling dataset for at least one of a geographical location and a time period; and
- providing, via the API, the dengue modelling dataset, optionally wherein the dengue modelling dataset is usable by a dengue outbreak prediction model to predict or estimate one or more parameters associated with and/or characteristic of a dengue outbreak.

[0011] For example, the dengue modelling dataset can be processible by an Artificial-Intelligence-based dengue outbreak prediction model (also referred to as “dengue prediction model” or similar herein) to predict one or more parameters associated with and/or characteristic of a dengue outbreak at said at least one of the requested geographical location and time period. Alternatively or additionally, the dengue modelling dataset may be processible by a time-series-based dengue prediction model.

[0012] By collecting raw dengue case data from a plurality of dengue case data sources and consolidating these data with raw weather data in the refined data repository, validity of the provided dengue modelling dataset can be ensured and the informative value or quality of the dengue modelling dataset can be significantly improved, for example when compared to current techniques applied to predict dengue outbreaks.

[0013] Also, the computer-implemented method of providing a dengue platform described herein can enable sophisticated computer-based surveillance of dengue disease. For instance, the method of the present disclosure can enable the implementation of reliable early warning systems that can reliably and consistently detect dengue outbreaks, from local to global scale. In turn, delays or inconsistencies in predictions or estimates for an upcoming dengue outbreak, for example in terms of its severeness, timing and/or course over time, can be reduced or avoided. Hence, aspects of the present invention will enable public health officials and policymakers to timely prepare for potential outbreaks and educate the public about prevention measures. In addition, the impact of outbreaks on the healthcare system can be minimized and patients can receive timely and appropriate care. Also, by predicting parameters associated with a dengue outbreak (synonymously or interchangeably used herein with precursors associated with a dengue outbreak), such as, e.g., the severity and timing of said outbreak, effective medical, manufacturing and commercial strategies can be developed for engagement, production, distribution and outreach of corresponding vaccines.

[0014] As used herein, “collecting raw data from a data source in a raw data repository,” or similar terms, in particular collecting raw dengue case data from a dengue case data source and raw weather data from a weather data source, can refer to or include obtaining and/or downloading the data from the respective data source in its raw data structure or format. Accordingly, raw data may refer to data being stored in their raw format or structure, for example as stored on the respective raw data source. In an example, the raw data repository may refer to or include a replicate collection of the raw data collected or obtained from one or more data sources, for example raw dengue case data sources or at least one weather data source. Therein, the raw data may be collected in the raw data repository without modifying or altering an informational content of and/or data values contained in the raw data obtained or received from one or more dengue case data sources and/or at least one weather data source.

[0015] For example, the raw dengue case data as used herein may include and/or contain one or more data values, which may be indicative of one or more dengue cases reported for a particular geographical location and/or time by one or more dengue case data sources. Alternatively or additionally, the raw weather data as used herein may include and/or contain one or more data values, which may be indicative of one or more weather parameters reported for a particular geographical location and/or time by one or more weather data sources. The raw dengue case data values and the raw weather data may be included or incorporated in the raw data repository with their original or unmodified informational content, as retrieved from the one or more dengue case data sources and the weather data source. It should be noted that in the context of the present disclosure, any reference to data can include a reference to one or more data values indicated by the respective data.

[0016] In the context of the present disclosure, generating the refined data repository based on consolidating the raw dengue case data and the raw weather data from the raw data repository may refer to or include aggregating and/or combining the raw data in the refined data repository.

Optionally, the collected raw dengue case data and/or the raw weather data may be converted from the raw data into refined data, for example refined dengue case data and/or refined weather data, which are consistent to each other, for example in terms of one or both geographical location and time. As will be discussed in more detail hereinbelow, one or more data processing or modification steps can be applied to consolidate the raw dengue case data and the raw weather data (which may commonly referred to as "raw data" herein) in the refined data repository. Further optionally, additional data or information may be obtained from the raw data and included in the refined data repository.

[0017] Alternatively or additionally, a data structure or format of the refined data repository may differ from the raw data structure, for example allowing for fast and reliable queries of data from the refined data repository. Accordingly, by consolidating the raw dengue case data and raw weather data in the refined data repository, an informational content of the data of the refined data repository can be enriched and/or further handling of the data to generate or obtain the dengue modelling dataset can be improved.

[0018] In the context of the present disclosure, the dengue platform for predicting one or more precursors associated with a dengue outbreak and/or for predicting a dengue outbreak can refer to a data-driven computing platform configured to receive one or more API calls via the at least one API interface and to provide, in response thereto, one or more dengue modelling dengue datasets, for example for a requested geographical location and/or time period. Such dengue platform can be implemented or embodied in the computing system described herein. Therefore, the terms "dengue platform" and "computing system" may be interchangeably or synonymously used herein.

[0019] The computing system may include one or more processors for data processing, for example for executing software instructions or a computer program instructing the computing system to carry out steps of the method described herein. Accordingly, any one or more steps of the method described herein can be carried out or performed by the computing system, respectively the dengue platform.

[0020] In an exemplary implementation, the computing system may include or be communicatively coupled to one or more data storage systems or devices, for example for storing the raw data repository, the refined data repository and optionally other data. The computing system according to the present disclosure can be embodied as a standalone computing device or can comprise a plurality of interconnected and communicatively coupled computing devices, for example in the form of a cloud computing network.

[0021] The "API interface" or similar terms may refer to a communication interface of the dengue platform and/or computing system configured to receive one or more API calls from one or more external or remote computing devices and to communicate a response. For example, the API interface may refer to or include a web-based API and the computing system or dengue platform can be embodied as one or more web servers, for example communicating with one or more remote or external computing devices based on HTTP or HTTPS (Hypertext Transfer Protocol (Secure)). Any other implementation of application programmatic interface, for example based on another public or proprietary protocols, can be used instead of or in addition to.

[0022] In an exemplary implementation, the API interface can include one or more API endpoints responsive at a particular interface or network address, for example web address. The API interface and/or one or more API endpoints may be configured to receive an API call in a particular or predefined format, for example defined according to API-specific rules, as will be discussed in more detail hereinbelow, in particular with reference to the Figures.

[0023] In an exemplary embodiment, collecting the raw dengue case data in the raw data repository can include retrieving and/or downloading the raw dengue case data from the plurality of different dengue case data sources, and storing the raw dengue case data in a data structure corresponding to a data structure of the respective dengue case data source. Alternatively or additionally, collecting the raw dengue case data can include accessing the raw dengue case data at one or more dengue case data sources and/or transmitting raw dengue case data from one or more dengue case data sources to the dengue platform or computing system. Any appropriate communication means or data connection between the one or more dengue case data sources and the dengue platform or computing system may be utilized for this purpose. For instance, the raw dengue case data may be transmitted to the dengue platform or computing system via an internet connection, via a cable connection, wirelessly, a BUS connection, a USB connection, or a combination thereof.

[0024] According to an embodiment, the plurality of dengue case data sources can include one or more web pages, one or more external APIs, and/or one or more external databases. For example, one or more web pages may be accessed and raw dengue case data may be obtained or retrieved therefrom based on web scraping.

[0025] As used herein, web scraping (also referred to as web harvesting or web data extraction) is a process of extracting data from websites. This can include various forms of data extraction. For example, text, numerical values or data in general that is depicted on a website can be extracted based on image analysis, optical character recognition or similar technique, and copied from the website. Alternatively or additionally, HTML code (or code in general) of a website may be analyzed to extract data from the website. Alternatively or additionally, one or more selectors, for instance

cascading style sheet selectors (CSS selectors), can be used to identify one or more elements, e.g. HTML elements, on a website and copy the corresponding data associated with or pointed to by the selector.

[0026] Alternatively or additionally, one or more external APIs and/or corresponding API endpoints
5 may be requested by the dengue platform or computing system to obtain or retrieve the raw dengue case data. For instance, the computing system may be configured to transmit an API call to one or more external APIs to receive raw dengue case data, for example raw dengue case data for a particular location and/or time period. Alternatively or additionally, one or more external databases may be communicatively coupled to the dengue platform or computing system, and may be
10 accessed by the computing system or dengue platform to obtain or retrieve the raw dengue case data, for example based on querying raw dengue case data from the one or more external databases.

[0027] In an example, the raw data repository includes raw dengue case data of a plurality of heterogenous dengue case data sources. In other words, the plurality of dengue case data sources
15 may be heterogenous, which may mean that they contain or provide heterogenous raw dengue case data. For instance, raw dengue case data of different dengue case data sources can differ in one or more of data structure, data format, temporal granularity, time information, spatial granularity, and geographical location.

[0028] In an exemplary embodiment, the raw data repository includes raw dengue case data
20 indicative of dengue cases reported in one or more countries comprising Brazil, Singapore, Malaysia, Indonesia, Sri Lanka, and Argentina. While these countries are currently affected by dengue and may thus be of particular interest for investigating current developments of dengue fever, the present disclosure is not limited to these countries. This is, raw dengue case data of any other country or combination of countries can be used instead of or in addition to, such as for
25 example raw dengue case data indicative of dengue cases reported in one or more of Singapore, Indonesia, Malaysia, Brazil, Colombia, Argentina, Thailand, Mexico, Philippines, India, Sri Lanka and Vietnam. Using raw dengue case data indicative of reported dengue cases in a plurality of different countries can allow for a plausibility check of raw dengue case obtained from the different sources and can allow to compute, model or estimate a geographical spread of a dengue outbreak with high
30 precision and accuracy.

[0029] Alternatively or additionally, the raw data repository includes raw dengue case data from a plurality of data sources associated with a single country. Accordingly, a plurality of different dengue case data sources reporting dengue case data for the same country may be accessed and corresponding raw dengue case data may be collected in the raw data repository. Using raw dengue
35 case data indicative of reported dengue cases in a single country but from different dengue data sources therein can allow for a plausibility check of raw dengue case obtained from the different sources. Also, statistical precision and hence informative value of the raw dengue case data collected in the raw data repository can be enhanced, when compared to using a single dengue data source only.

[0030] In an exemplary embodiment, the method further comprises retaining, in the raw data repository, a copy of the raw dengue case data retrieved from the plurality of dengue case data sources. In other words, the raw data repository may contain a copy of the raw dengue case data retrieved or obtained from one or more dengue case data sources. Storing a copy of the raw dengue case data can increase flexibility of the computing system or dengue platform, for example to implement future requirements or take into consideration future developments or changes in the raw dengue case data reported by the one or more dengue case data sources.

[0031] In yet another exemplary implementation, the method further comprises re-retrieving raw dengue case data from at least one of said plurality of dengue case data sources, and determining one or more changes in dengue cases reported for a particular geographical location and/or a particular time period based on inter-comparing the re-retrieved raw dengue case data and the raw dengue case data collected in the raw data repository from said at least one dengue case data source. For example, raw dengue case data for a particular location and/or time period may be retrieved from one or more dengue case data sources at a first time instant, and the raw dengue case data for the same location and/or time period may be re-retrieved from said one or more dengue case data sources at a second time instant. This allows one to determine whether changes in the raw dengue case data have been recorded in said one or more dengue case data sources. Also, validity and/or plausibility of the raw dengue case data retrieved at different time instants from the same dengue case data source can be checked or confirmed.

[0032] In general, raw dengue case retrieved from one or more dengue case data sources and/or collected in the raw data repository is indicative of one or more dengue cases reported at one or more geographical locations for one or more time periods. For example, the raw dengue case data can be indicative of one or more of a number of dengue cases, a number of severe dengue cases, a number of laboratory confirmed dengue cases, or a number of dengue-related deaths. Alternatively or additionally, the raw dengue case data may be indicative or contain information about a country, a region, a sub-region, a state, and a city, for which the one or more dengue cases have been reported. Alternatively or additionally, the raw dengue case data may also contain demographic information, such as a size of a population, and/or a dengue serotype indicator indicative of a serotype of the dengue virus that caused one or more dengue infections. It is emphasized that also other information may be contained in the raw dengue case data.

[0033] In an exemplary embodiment, collecting the raw weather data in the raw data repository can include retrieving and/or downloading the raw weather data from at least one weather data source, and storing the raw weather data in a data structure corresponding to a data structure of the respective weather data source. Alternatively or additionally, collecting the raw weather data in the raw data repository can include accessing the raw weather data at the one or more weather data sources, and/or transmitting the raw weather data from the at least one weather data source to the computing system and/or dengue platform. Any appropriate communication means or data connection between the one or more weather data sources and the dengue platform or computing system may be utilized for this purpose. For instance, the raw weather data may be transmitted to

the dengue platform or computing system via an internet connection, via a cable connection, wirelessly, or a combination thereof.

[0034] According to an embodiment, the at least one weather data source can include one or more web pages, one or more external APIs, and/or one or more external databases. For example, one or more web pages may be accessed and raw weather data may be obtained or retrieved therefrom based on web scraping. Alternatively or additionally, one or more external APIs and/or corresponding API endpoints may be requested by the dengue platform or computing system to obtain or retrieve the raw weather data. For instance, the computing system may be configured to transmit an API call to one or more external APIs to receive raw weather data, for example raw weather data for a particular location and/or time period. Alternatively or additionally, one or more external databases may be communicatively coupled to the dengue platform or computing system, and may be accessed by the computing system or dengue platform to obtain or retrieve the raw weather data, for example based on querying raw weather data from the one or more external databases.

[0035] In an exemplary embodiment, the method further comprises retaining, in the raw data repository, a copy of the raw weather data retrieved from said one or more weather data sources. In other words, the raw data repository may contain a copy of the raw weather data retrieved or obtained from one or more weather data sources. Storing a copy of the raw weather data can increase flexibility of the computing system or dengue platform, for example to implement future requirements or take into consideration future developments or changes in the raw weather data reported by the one or more weather data sources.

[0036] In yet another exemplary implementation, the method further comprises re-retrieving raw weather data from the at least one weather data source, and determining one or more changes in the raw weather data reported for a particular geographical location and/or a particular time period based on inter-comparing the re-retrieved raw weather data and the raw weather data collected in the raw data repository from said at least one weather data source. For example, raw weather data for a particular location and/or time period may be retrieved from one or more weather data sources at a first time instant, and the raw weather data for the same location and/or time period may be re-retrieved from said one or more weather data sources at a second time instant later than the first time instant. This allows one to determine whether changes in the raw weather data have been recorded in said one or more weather data sources. Also, validity and/or plausibility of the raw weather data retrieved at different time instants from the same weather data source can be checked or confirmed.

[0037] According to an embodiment, collecting the raw dengue case data from the raw data repository in the refined data repository includes converting and/or transforming the raw dengue case data into refined dengue case data based on or by normalizing the raw dengue case data from one or more of the plurality of dengue case data sources. In other words, raw dengue case data indicative of dengue cases reported by the dengue case sources may be normalized, thereby generating refined dengue case data. It should be noted that a normalization of dengue case data can include or refer to a normalization of the data values, respectively, of the one or more dengue

cases reported at one or more geographical locations for one or more time periods, which the raw dengue case data is indicative of. Generally, normalizing the raw dengue case data can provide or ensure a consistent set of refined dengue case data in the refined data repository, for example consistent in terms of one or more normalization aspects or criteria applied.

5 [0038] In an example, converting the raw dengue case data into refined dengue case data can include normalizing the raw dengue case data with respect to one or more normalization aspects or criteria, for example by applying one or more normalization operations to the raw data values. The one or more normalization aspects or operations may be commonly applied to the raw dengue case data collected from the one or more dengue case data sources. Alternatively or additionally,
10 individual or different normalization operations may be applied to raw dengue case data obtained from one or more individual dengue case sources. Accordingly, the normalization applied to raw dengue case data, and optionally raw weather data, stored in the raw data repository can be modular. Therefore, new dengue case data sources, for example providing dengue case numbers from a new country, can easily be added to the dengue platform and the raw dengue case data
15 retrieved therefrom can be populated as refined dengue case data into the refined data repository. Likewise, existing dengue case data sources can easily be modified, exchanged and/or removed from the dengue platform. Also, weather data sources may be added, modified, exchanged and/or removed from the dengue platform. Hence, a highly modular dengue platform can be provided.

[0039] Optionally, the refined dengue case data in the refined data repository may be re-structured
20 with respect to the raw data repository and/or may be stored in a data format different than a data format in the raw data repository. This may allow to structure the refined data in the refined data repository in compliance with the API of the dengue platform, for example, such that the refined dengue case data can be efficiently consumed or requested by the API of the dengue platform.

[0040] Normalization aspects or criteria applied for normalizing the raw dengue case data may, for
25 example, be geographical location-based and/or time-based. In other words, a geographical location-based and/or time-based normalization may be applied to the raw dengue case data to generate the refined dengue case data. Optionally, other normalization operations, for example associated with other normalization aspects or criteria may be applied. For instance, a normalization of dengue case data in terms of one or more threshold or reference case numbers may be applied,
30 including a threshold case number of total dengue cases reported, a threshold case number of severe dengue cases, a threshold case number of laboratory confirmed dengue cases, and/or a threshold case number of dengue-related deaths.

[0041] Geographical location-based normalization can be synonymously used herein with a
normalization of the raw dengue case data in terms of geographical location, and time-based
35 normalization can be synonymously used with a normalization of the raw dengue case data in terms of time. For example, the raw dengue case data may be normalized in terms of a geographical location, such that the refined dengue case data is indicative of dengue cases reported for a common geographical location and/or such that the refined dengue case data match each other with respect to a geographical location. Alternatively or additionally, the raw dengue case data may be
40 normalized in terms of a geographical location, such that the refined dengue case data match each

other with respect to a spatial resolution, a spatial granularity and/or a spatial scale of reported dengue cases.

[0042] Alternatively or additionally, the raw dengue case data may be normalized in terms of time, such that the refined dengue case data match each other with respect to one or more of a time scale, a time format, a time zone, a time resolution and a temporal granularity of the reported dengue cases.

[0043] As used herein, a “temporal granularity” of dengue case data, or similarly used term(s), may denote a temporal resolution of the data and/or the dengue case numbers indicated by the dengue case data. For instance, a temporal granularity may be defined or given by the temporal distance between consecutively reported dengue case numbers. Likewise, a “spatial granularity” of the dengue case data, or similarly used term(s), may denote a spatial resolution of the data and/or the dengue case numbers indicated by the dengue case data. For instance, a spatial granularity may be defined or given by the spatial distance between locations, for which dengue cases have been reported.

[0044] Normalizing the raw dengue case data with respect to location and/or time, and optionally other aspects, can increase inter-comparability of the normalized dengue case data, respectively, the refined dengue case data. Also, by normalizing the raw dengue case data, meaningful aggregation and/or grouping of the refined dengue case data may be enabled, which can further increase statistical significance of the refined dengue case data, and hence also of the dengue modelling dataset. Moreover, normalizing the raw dengue case data of different dengue case data sources can allow to add additional dengue case data sources as well as remove or replace existing dengue case data sources considered in or coupled to the dengue platform. Hence, a modular dengue platform can be provided, allowing to modularly change, modify, supplement and/or expand the dengue case data sources considered and/or communicatively coupled thereto.

[0045] In an exemplary embodiment, normalizing the raw dengue case data in terms of geographical location includes determining a number of dengue cases reported for a particular geographical location with one or more predefined spatial granularities. A spatial granularity may, for example, be defined by or associated with a particular spatial expansion of a region or area, in which dengue case numbers may be counted or added to provide a single dengue case number for said geographical location and with said spatial granularity. Alternatively or additionally, ranges of geographical coordinates, tiles of particular size or area and/or other measures can be utilized to define a particular spatial granularity for normalizing the raw dengue case data.

[0046] In an exemplary implementation, said one or more spatial granularities can be associated with one or more of a country, a region in a country, and a sub-region in a country. Accordingly, the raw dengue case data may be normalized, such that they are indicative or representative of dengue cases reported in a particular country, region and/or sub-region. Therein, regions or sub-regions may be defined according to various different space-related criteria. For example, states of a country may be used to define regions of a country. Alternatively or additionally, postal codes, counties and/or cities in a country may be used to define sub-regions of a country.

[0047] According to an embodiment, “normalizing the raw dengue case data in terms of geographical location,” or similarly used term(s), includes determining and/or computing, by means of the computing system, a number of dengue cases reported for a particular geographical location with a plurality of different predefined spatial granularities. For example, based on the raw dengue case data, refined dengue case data may be computed with different spatial granularities, such as dengue case data or numbers for countries, regions, and sub-regions. This may increase versatility and flexibility of the dengue platform and/or refined data repository, for example allowing to efficiently generate or provide a requested dengue modelling dengue set in response to an API call.

[0048] According to an embodiment, normalizing the raw dengue case data in terms of time includes determining a number of dengue cases reported for a particular period of time with one or more predefined temporal granularities. Exemplary and non-limiting temporal granularities can be associated with or be defined by one or more of a day, a week, a month, a plurality of months, and a year. However, any other temporal granularity or temporal resolution of the dengue case data may be used instead of or in addition to.

[0049] In an exemplary implementation, “normalizing the raw dengue case data in terms of time,” or similarly used term(s), includes determining and/or computing, by means of the computing device, a number of dengue cases reported for a particular geographical location with a plurality of different predefined temporal granularities. For example, based on the raw dengue case data, refined dengue case data may be computed with different temporal granularities, such as dengue case data or numbers for one or more of a day, a week, a month, a plurality of months, and a year. This may increase versatility and flexibility of the dengue platform and/or refined data repository, for example allowing to efficiently generate or provide a requested dengue modelling dengue set in response to an API call.

[0050] For example, the computing device may be configured to compute refined dengue case data indicative of dengue cases reported at one or more geographical locations in one or more of a daily, weekly, a monthly and a yearly format. Optionally, refined dengue case data indicative of dengue cases reported at one or more geographical locations with one or more spatial granularities may be computed and stored in the refined data repository.

[0051] In an example, the refined data repository may include refined dengue case data indicative of one or more of a number of dengue cases, a number of severe dengue cases, a number of laboratory confirmed dengue cases, a number of dengue-related deaths, a country, a region, a sub-region, a population, and a dengue serotype indicator. Generally, the information contained in the raw dengue case data and obtainable from the one or more dengue case data sources may be completely contained in the refined dengue case data, thereby ensuring that all available information on dengue cases reported by the dengue case data sources is retained in the refined data repository.

[0052] In yet another exemplary implementation, collecting the raw weather data from the raw data repository in the refined data repository includes converting the raw weather data into refined weather data based on normalizing the raw weather data in terms of one or both geographical location and time. Depending on the availability of raw weather data from the at least one weather

data source, the normalized or refined weather data may be retrieved directly from the weather data source, for example based on transmitting a corresponding API call to the weather data source.

Optionally, however, the raw weather data may be retrieved and processed by the dengue platform or computing system to compute or generate the refined weather data stored in the refined data

5 repository. Generally, similar or identical normalization operations, aspects or criteria as those applied to raw dengue case data can be considered for normalizing the raw weather data, thereby ensuring consistency of the refined weather and dengue case data in terms of one or more normalization aspects, in particular in terms of one or both geographical location and time.

[0053] The raw weather data can be indicative of one or more weather parameters including at least
10 one of temperature, humidity, and precipitation. Optionally, any other weather- or climate-related parameter obtainable from one or more weather data sources can be included in the raw weather data.

[0054] In an example, collecting the raw weather data from the raw data repository into the refined data repository may include determining refined weather data based on the raw weather data, the
15 refined weather data being indicative of one or more of a mean value, a maximum value, and a minimum value of one or more weather parameters with one or more predefined spatial granularities and/or one or more predefined temporal granularities. In other words, one or more of a mean value, a maximum value, and a minimum value of one or more weather parameters with one or more predefined spatial granularities and/or one or more predefined temporal granularities may be
20 computed by the computing system and stored in the refined data repository.

[0055] For example, the raw weather data may be grouped and/or aggregated according to the one or more predefined spatial granularities and/or the one or more predefined temporal granularities, and stored as refined weather data in the refined data repository.

[0056] Generally, the weather parameters indicated by the refined weather data may be particularly
25 useful for many dengue prediction models, which may utilize these parameters as predictors in the corresponding model. This, however, applies to any refined data stored in the refined data repository, including the refined dengue case data and the refined weather data.

[0057] Accordingly, the refined data stored in the refined data repository may be representative of one or more predictors usable by various prediction models and obtainable by requesting a
30 corresponding dengue modelling dataset via the API of the dengue platform. Multiple different dengue prediction models, including models based on historic-analysis and machine learning based models, can thus request accurate data from the dengue platform via API calls to the API of the dengue platform, which can be used as predictors for the computation of one or more precursors associated with a dengue outbreak or dengue infection in general. Also, reliable predictions can be
35 provided in timely manner or even in real-time based on the available data, which can enable implementation of efficient early warning systems at the affected locations or areas to minimize a potential effect of a predicted dengue outbreak.

[0058] According to an embodiment collecting the raw weather data from the raw data repository into the refined data repository includes deriving and/or computing, by means of the computing
40 device, one or more weather-related parameters from the raw weather data, the weather-related

parameters including at least one of an amount of rainfall at a geographical location within a predetermined period of time, a minimum predetermined number of consecutive days with rainfall, and a minimum predetermined number of consecutive days with a predetermined amount of rainfall per day. The method may further comprise storing the computed one or more weather-related parameters in the refined data repository. Hence, additional information may be derived from the raw weather data and stored in the refined data repository, thereby increasing an informative value of the refined data repository with respect to the raw data repository. Also, the weather-related parameters may be included in a dengue modelling dataset in response to a corresponding API call received from one or more external computing devices, which may be supplied to one or more dengue outbreak prediction models implemented on the external computing device.

[0059] According to an embodiment, collecting the raw dengue case data and the raw weather data from the raw data repository in the refined data repository includes analyzing the raw dengue case data and/or the raw weather data in terms of missing data values, and compensating for missing data values in the raw dengue case data and/or raw weather data based on data imputation.

Therein, compensating for missing data values may include calculating and/or generating new data values for the dengue case data and/or weather data, which new data values may be included as refined dengue case data and/or refined weather data in the refined data repository.

[0060] For example, one or more new dengue case data values may be generated and/or computed based on interpolating and/or extrapolating two or more raw dengue case data values indicated by or contained in the raw data repository. Alternatively or additionally, one or more new weather data values may be generated and/or computed based on interpolating and/or extrapolating two or more raw weather data values indicated by or contained in the raw data repository. Missing data values or missing data in one or both the raw dengue case data and the raw weather data, if used as input for a dengue prediction model, could adversely affect an accuracy and robustness of a prediction provided by the respective dengue outbreak prediction model. Hence, by compensating for missing values, a complete set of data and data values may be provided in the refined data repository, based on which the dengue platform or computing system can generate a complete dengue modelling dataset that can be supplied to a dengue outbreak prediction model to provide a prediction for a dengue outbreak with high accuracy and reliability.

[0061] According to an embodiment, collecting the raw dengue case data and the raw weather data from the raw data repository into the refined data repository includes analyzing the raw dengue case data and/or the raw weather data for outliers, and removing one or more outliers from the raw dengue case data and/or the raw weather data. For example, outliers may result from erroneous measurements and/or data recordings of raw dengue case data and/or raw weather data, which may adversely affect a prediction based on such data. Hence, by removing such outliers, an overall data quality of the refined data in the refined data repository can be improved, which can enable an accurate prediction for a dengue outbreak by a dengue outbreak prediction model.

[0062] According to an embodiment, generating the dengue modelling dataset based on the refined dengue case and/or weather data includes analyzing the refined dengue case data and/or the refined weather data for outliers, and removing one or more outliers from the refined dengue case

data and/or the refined weather data. For example, outliers may result from erroneous measurements and/or data recordings of (raw) dengue case data and/or (raw) weather data, which may adversely affect a prediction based on such data. Hence, by removing such outliers, an overall data quality of the dengue modelling dataset can be improved, which can enable an accurate prediction for a dengue outbreak by a dengue outbreak prediction model.

[0063] Optionally, the refined data or data values in the refined data repository, including the refined dengue case data and the refined weather data, can be further processed by the dengue platform based on applying one or more mathematical operations to the data or data values. For example, the refined data may be smoothed, e.g. based on applying a running average or other mathematical operation, the refined data may be scaled by an offset, shifted in time, and/or shifted in location. Any other mathematical operation can in principle be applied to the refined data to either modify or update the refined data or generate new refined data or data values that can be included in the refined data repository.

[0064] According to an embodiment, the API call, received by the computing system or dengue platform, includes a request from an external computing device to obtain a dengue modelling dataset for a requested geographical location and/or with a requested spatial granularity. In other words, the API call may define the requested geographical location and/or spatial granularity, for which a dengue modelling dataset is to be obtained or returned by the dengue platform. Hence, external computing devices and/or their users can request and receive a dengue modelling dataset in accordance with their interest and needs.

[0065] The requested geographical location and/or spatial granularity may be encoded or included in the API call. For example, one or more parameters indicating the requested geographical location and/or spatial granularity may be included in the API call, which can be processed or interpreted by the API of the dengue platform and trigger generation of the dengue modelling dataset, for example by triggering a corresponding API function call. It should be noted that the one or more parameters included in the API call can define one or both a geographical location and a spatial granularity, which allows to obtain a dengue modelling dataset based on one or both the geographical location and a spatial granularity. Alternatively or additionally, a dedicated API endpoint may be provided by the API of dengue platform, which when requested with a generic or predefined API call, returns the dengue modelling dataset for a particular geographical location and/or spatial granularity.

[0066] Optionally, an external computing device may authenticate at the API of the dengue platform to obtain a dengue modelling dataset. For example, an API token and API secret, pre-shared among the external computing device and the dengue platform, may be used to authenticate the external computing device. Alternatively or additionally, other authentication means, for example based on user credentials or an authentication certificate shared among the external computing device and the dengue platform can be used for authentication purposes. Optionally, a dedicated API endpoint may be provided for authenticating at the API of the dengue platform.

[0067] An authentication of the external computing device at the API of the dengue platform may be performed prior to requesting a dengue modelling dataset. For example, a first API call may serve

for authentication and a second API call may request the dengue modelling dataset. Alternatively, a single API call may serve for both authentication and requesting the dengue modelling dataset.

[0068] According to an embodiment, the method further comprises, in response to receiving a request from an external computing device to obtain a dengue modelling dataset for a requested geographical location and/or with a requested spatial granularity, retrieving refined dengue case data and/or refined weather data from the refined data repository for the requested geographical location and/or with the requested spatial granularity, and providing the dengue modelling dataset based on the refined dengue case data and/or the refined weather data. Accordingly, a dengue modelling dataset tailored around the specific needs and demands of a user of an external computing device can be generated and provided by the dengue platform, for example which can then be used by dengue outbreak prediction models. Hence, many different dengue prediction models can be fed with dengue modelling datasets provided by the dengue platform.

[0069] According to yet another embodiment, the received API call includes a request from an external computing device to obtain a dengue modelling dataset for a particular country, region or sub-region, and wherein the method further comprises retrieving refined dengue case data and/or refined weather data from the refined data repository for the requested country, region or sub-region, and providing the dengue modelling dataset based on the refined dengue case data and/or the refined weather data. In other words, the API call may define one or more countries, regions and/or sub-regions, for which a dengue modelling dataset is to be obtained or returned by the dengue platform. Hence, external computing devices and/or their users can request and receive a dengue modelling dataset in accordance with their interest and needs.

[0070] The API call may, for example, contain information about the one or more countries, regions and/or sub-regions. For example, the names of one or more countries or one or more country codes may be included in the API call to request a dengue modelling dataset for said one or more countries. Alternatively or additionally, one or more identifiers of one or more regions and/or sub-regions may be contained in the API call, such as for example one or more names of regions or sub-regions, state names or state codes identifying states of a country, one or more postal codes of particular regions or sub-regions, geographical coordinates and others. Alternatively or additionally, a dedicated API endpoint may be provided by the API of the dengue platform, which when requested with a generic or predefined API call, returns the dengue modelling dataset for a particular country, region and/or sub-region.

[0071] According to an embodiment, the received API call includes a request from an external computing device to obtain a dengue modelling dataset for a requested time period and/or with a requested temporal granularity. In other words, the API call may define the requested time period and/or temporal granularity, for which a dengue modelling dataset is to be obtained or returned by the dengue platform. Hence, external computing devices and/or their users can request and receive a dengue modelling dataset in accordance with their interest and needs.

[0072] The requested time period and/or temporal granularity may be encoded or included in the API call. For example, one or more parameters indicating the requested time period and/or temporal granularity may be included in the API call, which can be processed or interpreted by the API of the

dengue platform and trigger generation of the dengue modelling dataset. For example, a date range and/or one or more identifiers for one or more temporal granularities, such as daily, weekly, monthly and/or yearly, can be included in the request. It should be noted that the one or more parameters included in the API call can define one or both a time period and a temporal granularity, which allows to obtain a dengue modelling dataset based on one or both the time period and the temporal granularity. Hence, for example all dengue data available can be requested in yearly or other granularity, or only data within a particular period of time may be requested. Alternatively or additionally, a dedicated API endpoint may be provided by the API of dengue platform, which when requested with a generic or predefined API call, returns the dengue modelling dataset for a particular time period and/or temporal granularity.

[0073] Also, it should be noted that time-based requests can be combined with location-based requests, or optionally requests based on other parameters, which allows to obtain dengue modelling datasets for a particular time period and/or location, optionally with a particular temporal and/or spatial granularity.

[0074] According to an embodiment, the method further comprises in response to receiving a request from an external computing device to obtain a dengue modelling dataset for a requested time period and/or with a requested temporal granularity, retrieving refined dengue case data and/or refined weather data from the refined data repository for the requested time period and/or with the requested temporal granularity, and providing the dengue modelling dataset based on the refined dengue case data and/or the refined weather data retrieved from the refined data repository.

[0075] According to an embodiment, generating and/or providing the dengue modelling dataset comprises selecting the refined dengue case data and/or refined weather data from the refined data repository. The selected refined dengue case and/or weather data may for example be aggregated and/or combined in the dengue modelling dataset. Alternatively or additionally, the selected refined dengue case and/or weather data may be converted into a predefined data structure or format, such as a human-readable or machine-readable format. In a non-limiting example, the requested dengue case and/or weather data may be converted into a text string containing and/or being indicative of the requested dengue case and/or weather data (or corresponding data values), which may be returned to the external computing device as dengue modelling dataset in response to an API call.

However, any other form of data output, for example in encrypted form or binary form, can be implemented.

[0076] According to an embodiment, the method further comprises converting the received API call into a database query for querying the requested dengue modelling dataset from the refined data repository. In other words, a database query may be generated by the dengue platform based on the received API call to select, query and/or retrieve the requested dengue modelling dataset, respectively, the requested dengue case and/or weather data.

[0077] Depending on the actual API call and data requested, the database query may be a location-based database query, a time-based database query, or a spatio-temporal database query.

Accordingly, refined dengue case and/or weather data may be queried from the refined data repository in accordance with the API call, based on which an external computing device may

request to obtain a dengue modelling dataset for a requested time period, a requested geographic location or both. Optionally, one or more spatial and/or temporal granularities may be defined by the API call and reflected or contained in the database query.

[0078] Optionally, refined dengue case and/or weather data may be further processed by the computing system and/or dengue platform to provide the dengue modelling dataset as requested. For instance, grouping, interpolation, extrapolation, aggregation or other operations may be applied to the refined data to generate the dengue modelling dataset.

[0079] Generally, the dengue modelling dataset generated by the dengue platform and provided to the external computing device in response to the API call can include any of the information, data, and/or data values contained in the refined data repository and/or derivable therefrom. Exemplary data, data values, or information in general, indicated by the dengue modelling dataset can include one or more of a number of severe dengue cases, a number of laboratory confirmed dengue cases, a number of dengue-related deaths, a population, a dengue case data source, a dengue serotype indicator, a data quality, a country, a region, a sub-region, a state, a temporal granularity, and a spatial granularity. Any one or more of the aforementioned data, data values or information may be specifically requested by the external computing device, for example by including corresponding parameters, identifiers or flags in the API call. This allows to increase flexibility of the dengue platform and allows to provide many different dengue modelling datasets for many different dengue outbreak prediction models, which may be based on different combinations of predictors.

[0080] According to a further embodiment, the dengue modelling dataset is indicative of one or more of a minimum temperature, a maximum temperature, an average temperature, a humidity, and a precipitation at the requested geographical location and/or for a requested period of time.

Alternatively or additionally, the dengue modelling dataset may be indicative of one or more of a country of a weather station, a measurement location, a county of a weather station, a municipality of a weather station, a city of a weather station, a postcode of a weather station, a latitude of a weather station, and a longitude of a weather station. Alternatively or additionally, the dengue modelling dataset may be indicative of one or more of an amount of rainfall, a minimum predetermined number of consecutive days with rainfall, and a minimum predetermined number of consecutive days with a predetermined amount of rainfall per day. Accordingly, any one or more of the aforementioned data, data values, parameters or information may be specifically requested by the external computing device, for example by including corresponding parameters, identifiers or flags in the API call. This allows to increase flexibility of the dengue platform and allows to provide many different dengue modelling datasets for many different dengue outbreak prediction models, which may be based on different combinations of predictors.

[0081] According to an embodiment, the method further comprises providing the dengue modelling dataset as input to an AI-based dengue outbreak prediction model to predict one or more precursors associated with a dengue outbreak. The AI-based outbreak prediction model may be implemented on the dengue platform or computing device. Alternatively, the AI-based dengue prediction model may be implemented on an external computing device communicatively coupled to the dengue platform. It should be noted that the present disclosure is not limited to AI-based dengue outbreak

prediction models, but dengue modelling datasets may also be provided to and processed by time-series-based dengue prediction models.

[0082] According to an embodiment, the method further comprises predicting, based on processing the dengue modelling dataset with a dengue outbreak prediction model, for example an AI-based dengue prediction model, one or more parameters associated with and/or characteristic of a dengue outbreak and/or characteristic of a dengue outbreak. For example, the dengue outbreak prediction model may be configured to compute one or more of a probability of an onset of a dengue outbreak, a time of an onset of a dengue outbreak, a geographical location of a dengue outbreak, an epicenter of a dengue outbreak, a number of expected cases of a dengue outbreak, a number of expected cases of a dengue outbreak within a predetermined period of time, and a severity measure for a dengue outbreak. This may allow for a reliable and accurate characterization of a predicted dengue outbreak, and may allow to estimate its potential impact.

[0083] Optionally, the method may comprise collecting further dengue-related and/or weather-related data from one or more further data sources including social media platforms, world bank data, open data, research data, scientific publications, publication data, social data, building density data, vector-related data, mosquito data. Any one or more of the aforementioned data, data values or information may be included in the refined data repository, and ingested by the dengue prediction model and used as predictor to compute one or more parameters associated with and/or characterizing a dengue outbreak.

Aspects related to the computer program and computing system

[0084] A further aspect of the present disclosure relates to a computer program, which, when executed by one or more processors of a computing system or dengue platform, instructs the computing system or dengue platform to perform steps of the method, as described herein.

Accordingly any step of the method described herein can be reflected in the computer program and/or be performed by the computing system or dengue platform.

[0085] A further aspect of the present disclosure relates to a computer-readable medium, for example a non-transitory computer-readable medium, storing a computer program, which, when executed by one or more processors of a computing system or dengue platform, instructs the computing system or dengue platform to perform steps of the method, as described herein.

[0086] Yet another aspect of the present disclosure relates to a computing system for predicting one or more precursors associated with a dengue outbreak, wherein the computing system includes processing circuitry with one or more processors and is configured to perform steps of the method as described herein.

[0087] The computing system may refer to or include the dengue platform. For example, the computing system may comprise a raw data repository storing the raw dengue case data from a plurality of different dengue case data sources, and storing raw weather data from one or more weather data sources. The computing system may further comprise an API configured to receive an API call from an external computing device and configured to provide a dengue modelling dataset in response to said API call.

[0088] Optionally, a dengue outbreak prediction model may be implemented in the computing system, for example an Artificial-Intelligence-based dengue outbreak prediction model configured to receive the dengue modelling dataset as input and to compute one or more parameters and/or precursors associated with a dengue outbreak based on processing the dengue modelling dataset.

5 BRIEF DESCRIPTION OF THE DRAWINGS

[0089] Exemplary embodiments will be further described with reference to Figures, wherein:

[0090] Figure 1 shows an exemplary computing system and dengue platform for predicting one or more parameters associated with a dengue outbreak;

10 [0091] Figure 2 shows a flowchart illustrating a computer-implemented method of providing or operating a dengue platform for predicting one or more parameters associated with a dengue outbreak; and

[0092] Figure 3 shows a block diagram illustrating operation of an exemplary dengue platform for predicting one or more parameters associated with a dengue outbreak.

15 [0093] The Figures are schematic only and not true to scale. In principle, identical or like parts, elements and/or steps are provided with identical or like reference numerals in the Figures.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0094] To protect public health, for example to prepare for potential dengue outbreaks and take corresponding countermeasures, numerous different dengue prediction or forecasting models may be developed, such as for use by numerous scientific working groups, agencies and institutions
20 worldwide. Such dengue prediction or forecasting models, also referred to herein as “dengue outbreak prediction models” or similar, may use data indicative of or at least related to the spread of dengue disease. Typical data that may be considered and used by dengue prediction models are climate or weather data and reported historic data on dengue cases, but also include other data sources, such as for example data related to the mosquito population, wherein the data relied upon
25 by the models are often referred to as “predictors” or precursors of the models or similar. However, different models for predicting dengue outbreaks may be based on different predictors or data considered for predicting a dengue outbreak and may rely on data from different data sources with strongly varying data quality, validity and availability. In addition, the dengue prediction models may differ in the actual modelling approach applied. For example, the majority of the dengue prediction
30 models may be based on statistical or time-series analysis of historic dengue case data or on machine learning techniques are utilized in current models.

[0095] These differences in dengue prediction models and dengue modelling approaches can lead to different predictions or estimates, wherein the differences in the quality and availability of the data utilised by the individual models may have the largest impact. For example, different models used in
35 conjunction with data of differing quality and availability can lead to diverse predictions of both the severity and the time-evolution or timing of a dengue outbreak at a given location. This can further lead to confusion and delays in taking appropriate countermeasures to minimize a potential dengue

outbreak. It may, therefore, be desirable to provide for an improved dengue data platform, also referred to herein as “dengue platform” or similar, for predicting a dengue outbreak or dengue infection in general, which can overcome or at least mitigate the aforementioned drawbacks used to predict dengue outbreaks.

5 [0096] Figure 1 shows an exemplary dengue platform 100 embodied or implemented in a computing system 101 for predicting one or more parameters and/or precursors associated with a dengue outbreak. The arrows in Figure 1 illustrate a data flow.

[0097] The dengue platform 100 or computing system 101 comprises a processing circuitry 102 with one or more processors 104 for data processing. The computing system 101 may be a single
10 computer, multiple computers, or a distributed system (e.g., a cloud environment) having scalable / elastic computing resources (e.g., data processing hardware) and/or storage resources (e.g., memory hardware). Data stores (e.g., a remote storage device) may be overlain on the storage resources to allow scalable use of the storage resources by a client and/or computing resources. The memory (e.g., the storage resources) stores information non-transitorily within the computing
15 system 101. The memory may be a computer-readable medium, a volatile memory unit(s), or non-volatile memory unit(s). The non-transitory memory may be physical devices used to store programs (e.g., sequences of instructions) or data (e.g., program state information) on a temporary or permanent basis for use by the computing device. The dengue platform 100 or computing system 101 further comprises a communication interface 110 communicatively couplable and/or coupled to
20 one or more data sources 200a, 200b, 200c, 200d (e.g., via a network). The communication interface 110 may be configured for wired communication, wireless communication or both. For example, an Internet, network, WiFi, LAN, USB, BUS, Bluetooth or other connection may be used to communicate with the one or more data sources 200a-200d and/or retrieve raw data therefrom.

[0098] The dengue platform 100 or computing system 101 further comprises a raw data repository
25 130 (e.g., at one or more data stores) storing raw data and a refined data repository 140 storing refined data. More specifically, the dengue platform 100 or computing system 101 of Figure 1 is communicatively coupled via the communication interface 110 to a plurality of dengue case data sources 200a, 200b, 200c, and to at least one weather data source 200d. However, any other number of dengue case data sources and weather data sources may be utilized. In particular, the
30 dengue platform 100 may be expanded modularly in terms of data sources connected thereto, as indicated by the three dots in Figure 1.

[0099] Exemplary dengue case data sources 200a, 200b, 200c and/or weather data sources may include one or more web pages, one or more external APIs, and one or more external databases, which can be communicatively coupled to the communication interface 110 to retrieve raw data
35 therefrom.

[0100] Exemplary dengue case data sources 200a-200c may report dengue cases from one or more countries, such as one or more of Colombia, Thailand, Mexico, Philippines, India, Vietnam, Brazil, Singapore, Malaysia, Indonesia, Sri Lanka, and Argentina or others. Raw dengue case data may, for example, be retrieved from ministries, institutions, universities, health organizations or
40 others. Generally, public and private dengue case sources can be used. For instance, raw dengue

case data may be retrieved and/or collected from Tycho, Ministry of Health Singapore, Ministry of Health Malaysia, Ministry of Health Indonesia, Ministry of Health Sri Lanka, Ministry of Health Argentina, FGV, World Health Organization, Pan American Health Organization, and others. An exemplary weather data source 200d may be the National Oceanic and Atmospheric Administration, NOAA, but other weather or climate services may be used instead or in addition.

[0101] The dengue platform 100 or computing system 101 is configured to retrieve and/or collect raw dengue case data from the plurality of dengue case data sources 200a-200c, and store a copy of the raw dengue case data in the raw data repository 130. The dengue platform 100 or computing system 101 is further configured to retrieve and/or collect raw weather data from the at least one weather data source 200d, and store a copy of the raw weather data in the raw data repository 130.

[0102] The raw dengue case data and/or raw weather data may, for example, refer to time-series data containing one or more data values representative of one or more parameters related to dengue infection, respectively related to weather or climate conditions.

[0103] In particular, the raw dengue case data is indicative of one or more dengue cases reported at one or more geographical locations for one or more time periods. For instance, the raw dengue case data may be indicative of one or more of a number of dengue cases, a number of severe dengue cases, a number of laboratory confirmed dengue cases, a number of dengue-related deaths, a country, a region, a sub-region, a state, a city, a population, and a dengue serotype indicator. Other information, data or parameters may be included in the raw dengue case data.

[0104] The raw weather data is indicative of one or more weather parameters at one or more geographical locations and one or more time periods. Exemplary weather parameters indicated by the raw weather data can include temperature, humidity, and precipitation, but any other parameter related to weather, climate and/or climate change can be contained in the raw weather data and stored in the raw data repository 130.

[0105] The general purpose of the raw data repository 130 may be to download raw dengue case data from various dengue case data sources 200a-200c and raw weather data from at least one weather data source 200d and store the raw data as copy in the same structure as the sources 200a-200d. This allows for a greater flexibility for future applications.

[0106] Also, raw dengue case data and/or raw weather data may be re-retrieved from the same data source 200a-200d, and analyzed by the computing system 101 to determine changes between the raw data retrieved at different times. Based on determined differences, raw dengue case data and/or raw weather data may be flagged as potentially erroneous, removed, or modified, for example.

[0107] Generally, raw dengue case data retrieved or collected from different dengue case data sources 200a-200c and/or raw weather data retrieved from one or more weather data sources 200d can differ in one or more aspects, such as data structure, data format, temporal granularity or resolution, time information, spatial granularity or resolution, geographical location, and other aspects.

[0108] To account for such heterogeneous raw dengue case data retrieved from one or more dengue case data sources 200a-200c, the raw dengue case data is converted into refined dengue

case data and stored in the refined data repository 140. Alternatively or additionally, the raw weather data is converted into refined weather data and stored in the refined data repository 140.

[0109] In particular, the raw dengue case data and/or raw weather data may be normalized by applying one or more normalization operations to the raw data, and/or raw data values indicated by or contained in the raw data. Since normalization may be applied to one or both the raw dengue case data and the raw weather data, it can be referred to hereinabove and hereinbelow to raw data, which includes one or both raw dengue case data and raw weather data.

[0110] For example, the one or more normalization operations may be applied commonly to all raw dengue case data and/or raw weather data retrieved from all sources 200a-200d. Alternatively or additionally, individual and/or different normalization operations may be applied to raw dengue case data retrieved from a particular source 200a-200c, and/or raw weather data retrieved from one or more weather data sources 200d.

[0111] The one more normalization operations applied to the raw data may normalize the raw data and/or raw data values in terms of one more normalization aspects or criteria. In particular, time-based normalization and/or geographical location-based normalization may be applied. In other words, the raw data can be normalized by the dengue platform 100 or computing system 101 in terms of one or both geographical location and time.

[0112] Normalizing the raw dengue case data in terms of geographical location can include determining a number of dengue cases reported for a particular geographical location and/or with one or more predefined spatial granularities, for example associated with one or more of a country, a region in a country, and a sub-region in a country. In particular, a number of dengue cases reported for a particular geographical location with a plurality of different predefined spatial granularities may be computed for raw data from one or more data sources 200a-200c.

[0113] Normalizing the raw dengue case data in terms of time can include determining a number of dengue cases reported for a particular period of time and/or with one or more predefined temporal granularities, for example associated with one or more of a day, a week, a month, a plurality of months, and a year. In particular, a number of dengue cases reported for a particular geographical location with a plurality of different predefined temporal granularities may be computed for raw data from one or more data sources 200a-200c.

[0114] The refined dengue case data stored in the refined data repository 140 may at least contain the information as the raw dengue case data. Optionally, the refined dengue case data may be enriched in informational content with respect to the raw dengue case data. For example, one or more dengue-related parameters may be computed based on the raw dengue case data and stored in the refined data repository 140. Non-limiting examples include an average, a mean, a maximum, a minimum dengue case number, and others, which may optionally be computed as total values, over a certain period of time, over a certain geographical location or area, or both of the latter.

[0115] Similar or identical normalization operations may be applied to the raw weather data as applied to the raw dengue case data. In particular, any normalization or normalization operation as well as any other data manipulation, modification and/or analysis described herein with reference to the raw dengue case data may be applied to the raw weather data, and vice versa.

[0116] The dengue platform 100 or computing system 101 may be configured to analyse raw data from different sources 200a-200d, inter-compare them and determine based on the result of the inter-comparison one or more normalization operations to be applied to the raw data of one or more data sources 200a-200d to compensate for differences in the raw data from different sources 200a-200d.

[0117] In general, the dengue platform 100 or computing system 101 is highly modular, so plugging in or connecting a new data source, for example for a new country, replacing and/or removing data sources can be quickly and efficiently be done, without interrupting the system 101.

[0118] Moreover, data analysis routines may be used to analyse the refined data or at least part thereof, which may be modified, removed and/or supplemented, without losing information on the original raw data as these are retained in the raw data repository. This provides for a highly flexible dengue platform 100 that can be adapted to future requirements.

[0119] For example, the raw dengue case data and/or raw weather data may be modified, updated and/or supplemented based on one or more modification operations or mathematical operations in general, which may be applied to one or both the raw data and the corresponding refined data obtained therefrom. Exemplary modification operations may include, data amputation, outlier detection, outlier removal, outlier flagging, re-scaling, smoothing, shifting, grouping, interpolation, extrapolation, fitting, and re-binning.

[0120] Alternatively or additionally, the dengue platform 100 or computing system 101 can be configured to derive one or more dengue-related parameters from the raw dengue case data in the refined data repository. Alternatively or additionally, the dengue platform 100 or computing system 101 can be configured to derive one or more weather-related parameters from the raw weather data and store them in the refined data repository. Such derived parameters can supplement the refined data and increase their informative value. Also, such derived parameters may be used as predictors by dengue prediction models.

[0121] Exemplary weather-related parameters derivable from the raw weather data include an amount of rainfall at a geographical location within a predetermined period of time, a minimum predetermined number of consecutive days with rainfall, and a minimum predetermined number of consecutive days with a predetermined amount of rainfall per day. For instance, rainy day events or days with precipitation > 0 mm, > 10 mm, > 20 mm, > 50 mm, > 100 mm or more, recorded for two or more, for 10 or more, 20 or more, and/or 50 or more days may be computed and stored in the refined data repository. However, any other weather-related parameter derivable from the raw weather data may be stored in the refined data repository 140.

[0122] The raw dengue case data and the raw weather data may be aggregated or combined in the refined data repository 140, as well as consolidated in terms of one or both geographical location and time, optionally computed for one or more temporal and/or spatial granularities, as described hereinabove. Hence, the refined data repository 140 may constitute a comprehensive data pool allowing to retrieve reliable and consistent data related to or indicative of dengue infection and/or its spread with respect to one or both time and geographic location.

[0123] The dengue platform 100 or computing system 101 further comprises an API 150, which may be implemented as API layer or software layer in a communication interface 160, which may optionally be combined with communication interface 110.

[0124] In a non-limiting example, the API 150 can be embodied or implemented as a web API, for example providing one or more API endpoints that may be accessible at one or more web or network addresses. Any other implementation of the API 150 is envisaged in the context of the present disclosure.

[0125] The dengue platform 100 or computing system 101 is configured to receive, via the API 150, one or more API calls from one or more external computing devices 300a, 300b, 300c, requesting one or more dengue modelling datasets, for example usable by one or more dengue outbreak prediction models 302a, 302b, 302c implemented on the one or more computing devices 300a-300c, for data visualization, monitoring, alerting or other purposes.

[0126] The API 150 may be configured to accept a variety of different requests, which can allow users to obtain or request dengue modelling datasets according to their specific needs and demands, for instance for data visualization at one or more of the external computing devices 300a-300c, and/or for supplying the dengue modelling dataset to a dengue outbreak prediction model 302a-302c.

[0127] Optionally, the API 150 may provide multiple API endpoints, which may be associated with particular network, web or API addresses and may be requested to provide particular or predefined data or a predefined response in response to receiving an corresponding API call at the respective API endpoint. In a non-limiting example, external computing devices 300a-300c may authenticate at a first API endpoint via an API call, for example to confirm that the respective computing device 300a-300c is registered at the dengue platform 100. Once authenticated, computing devices 300a-300c may call a second API endpoint to obtain the dengue modelling dataset in return. Optionally, for example weather data and dengue case data could be provided at separate API endpoints. It should be noted, however, that multiple API endpoints are optional only and the API 150 can comprise a single API endpoint only.

[0128] External computing devices 300a-300c and/or users thereof can request dengue modelling datasets tailored around their specific needs. For this purpose, the API 150 can be configured to accept, process and interpret multiple different API calls, which may optionally include one or parameters reflecting the data requested by the respective users or external computing devices 300a-300c. For example, different API function calls, each associated with a particular API call of predefined schema and/or including predefined information, may be implemented. By providing the correct API call in accordance with one of the predefined schemas, the API 150 provides the correct or requested dengue modelling dataset in response, which contains the requested data.

[0129] Exemplary and non-limiting API function calls, which can be configured in the API 150 and triggered by providing the respective API call to the API 150, can include one or more of a request for retrieving dengue case data only, respectively, a dengue modelling dataset only containing dengue case data. Alternatively or additionally, one or more API function calls may allow retrieving weather data only, and/or a combination of weather and dengue case data. Alternatively or

additionally, one or more API function calls may allow to retrieve dengue case and/or weather by country, by region, by sub-region, by date, by date range, by epidemiological week, by temporal granularity, by spatial granularity, by data source, by data quality, by population, by severeness of dengue cases, by dengue serotype, by number of deaths, by number of laboratory confirmed dengue cases, or any combination thereof. In general, any information contained in the refined dengue case data and/or any attribute of the refined data can be encoded in a particular API function call. Alternatively or additionally, one or more parameters may be included in an API call, which can further define the dengue modelling dataset requested by the respective computing device 300a-300c or user.

[0130] Exemplary and non-limiting parameters, which may be included in one or more API calls associated with one or more API function calls programmed into the API 150, can include one or more of a number of severe dengue cases, a number of laboratory confirmed dengue cases, a number of dengue-related deaths, a population, a dengue case data source, a dengue serotype indicator, a data quality, a country, a region, a sub-region, a state, a temporal granularity, and a spatial granularity. Any one or more of these parameters can be combined with any one or more API function calls, which allows users to obtain dengue modelling datasets meeting their demands, while also ensuring consistency and high quality of the data.

[0131] Alternatively or additionally, additional criteria or parameters for the dengue modelling dataset may be specified in an API call. For example, the maximum number of data values contained in the dengue modelling dataset could be specified, only data of particular quality, and/or only data from a particular data source 200a-200d may be requested by defining corresponding parameters in the API call. As mentioned above with respect to the API function calls, in general, any information contained in the refined dengue case data and/or any attribute of the refined data can be included as parameter in an API call to obtain particular data, and/or to restrict or define the dengue modelling dataset according to one or more parameters.

[0132] The one or more parameters may be included as string, as integers, as numbers, as dates and/or as flags in the respective API call. To give an example, API calls may request only severe dengue cases for a particular country, or by country, or by date, or by date range, or by any other information or attribute of the refined data in the refined data repository 140.

[0133] A received API call, optionally with one or more parameters, can for example be converted by the dengue platform 100 or computing system 101 into a database query for querying, selecting and/or retrieving the refined data as requested by the computing device 300a-300c from the refined data repository 140. The computing system 101 or dengue platform 100 is further configured to generate the dengue modelling dataset based on the selected refined data, and to transmit the dengue modelling dataset to the computing device 300a-300c in return.

[0134] Accordingly, based on the respective API call transmitted to the API 150 and optionally one or more parameters included therein, the dengue platform 100 selects or filters the requested data from the refined data repository 140, for example based on one or more database queries, and provides the dengue modelling dataset in predefined format and structure in return.

[0135] For instance, data or data values contained in the dengue modelling dataset can be provided as text strings with appropriate delimiters, delimiting single data values or data attributes.

Alternatively or additionally, lists of data values, arrays of data values, hierarchically structured data values, such as in JSON format, or any other data format or structure can in principle be provided by the API 150. Alternatively or additionally, the dengue modelling dataset may be provided in binary format and/or encrypted.

[0136] As for dengue case data, also one or more dedicated API function calls and/or one or more parameters defining the requested weather data, can be considered in the API 150, such that weather data meeting the users' interests can be obtained. For example, dengue modelling datasets indicative of one or more of a minimum temperature, a maximum temperature, an average temperature, a humidity, and a precipitation, for example at a requested geographical location and/or for a requested period of time, optionally with one or more spatial and/or temporal granularities, or any combination thereof can be requested by corresponding API calls. Alternatively or additionally, dengue modelling datasets indicative of one or more of one or more of a country of a weather station, a measurement location, a county of a weather station, a municipality of a weather station, a city of a weather station, a postcode of a weather station, a latitude of a weather station, a longitude of a weather station an amount of rainfall, a minimum predetermined number of consecutive days with rainfall, and a minimum predetermined number of consecutive days with a predetermined amount of rainfall per day, or any combination thereof can be requested.

[0137] In the following non-limiting examples of API calls associated with respective API function calls of the API, corresponding parameters and parameter types in brackets are given:

[0138] An API call `getCasesByCountry(country: String)`, which includes a country identifier as string, returns dengue cases for the specified country. An API call `getCasesByCountryAndRegion(country: String, region: String)`, which includes a country and region identifier as string, returns dengue cases for the specified region in the specified country. An API call

`getCasesByCountryAndRegionAndDateRange(country: String, region: String, start: Date, end: Date)`, which includes a country and region identifier as string as well as a start and end date, returns dengue cases for the specified region in the specified country for date period of time between start and end date. Another API `getWeatherByDateRange (country: String, start: Date, end: Date)` which includes a country as string, a start and an end date, returns weather data for the specified date range. Following this schema, multiple different API calls or corresponding API function calls can be implemented in the API 150, which provides a highly flexible dengue platform that can serve dengue modelling datasets for different purposes, different dengue outbreak prediction models, for example using different predictors, and to multiple users worldwide.

[0139] Also, it is noted that one or more query languages may be used to retrieve data from the refined data repository 150. Exemplary and non-limiting query languages are the GraphQL, FLUX, PromQL, InfluxQL or any other suitable query language for the type of database used in the refined data repository 140.

[0140] The dengue platform 100 or computing system 101 is further configured to provide, via the API 150, the dengue modelling dataset, wherein the dengue modelling dataset is processible by a

dengue outbreak prediction model 302a-30c, in particular an Artificial-Intelligence-based dengue outbreak prediction model to predict one or more precursors associated with a dengue outbreak. The data or data values contained in the dengue modelling dataset may be used as predictors for the dengue prediction model 302a-30c to predict the one or more precursors associated with a
5 dengue outbreak, or associated with dengue infection in general.

[0141] Optionally, also the dengue platform 100 or computing system 101 may include a dengue outbreak prediction model 180, for example an AI-based dengue outbreak prediction model 180. It is noted that also the dengue platform 100 itself can request a dengue modelling dataset via the API 150 and provide it to the dengue prediction model 180.

10 [0142] The dengue platform 100, the computing system 101, the dengue prediction model 180 of the dengue platform, the external computing devices 300a-300c and/or the dengue outbreak prediction models 302a-302c implemented thereon can be configured to compute, based on the dengue modelling dataset, one or more of a probability of an onset of a dengue outbreak, a time of an onset of a dengue outbreak, a geographical location of a dengue outbreak, an epicenter of a
15 dengue outbreak, a number of expected cases of a dengue outbreak, a number of expected cases of a dengue outbreak within a predetermined period of time, and a severity measure for a dengue outbreak.

[0143] The dengue platform 100 may generate and/or train one or more dengue outbreak prediction models 302a-30c using data from the refined data repository 150. Optionally, the dengue outbreak
20 prediction models 302a-302c include one or more neural networks trained on training samples generated from the data of the refined data repository 150 (i.e., any combination of the weather data and the dengue case data). For example, the dengue platform generates training samples from the refined data repository 150 and trains the dengue outbreak prediction models 302a-302c using the training samples and a loss function (e.g., a regression loss function, such as a mean square error
25 loss function, a mean absolute error function, etc.). In some examples, the dengue platform 100 generates and/or trains the dengue outbreak prediction models 302a-302c in response to an API request (e.g., from a user of the dengue platform 100) via the API 150. The API request may specify one or more training parameters for training the dengue outbreak prediction models 302a-302c based on the data of the refined data repository and/or the specific probabilities to predict, as
30 discussed above. The dengue platform 100 may provide one or more predictions from the dengue outbreak prediction models 302a-302c via the API 150.

[0144] Optionally, other additional data, which may be weather- and/or dengue-related, such as data from social media platforms, world bank data, open data, research data, scientific publications, publication data, social data, building density data, vector-related data, mosquito data, or the like,
35 can be included in the refined data repository 140, provided by the API 150 in response to an API call, and/or used as predictors by one or more dengue outbreak prediction models 180, 302a-302c to compute one or more precursors associated with a dengue outbreak.

[0145] Figure 2 shows a flowchart illustrating a computer-implemented method of providing or operating a dengue platform or computing system, for example the dengue platform 100 or
40 computing system 101 described with reference to Figure 1, for predicting one or more parameters

and/or precursors associated with a dengue outbreak. Accordingly, any of the steps of the method can be performed by the dengue platform 100 or computing system 101 described with reference to Figure 1.

5 [0146] The method comprises, in step S1, collecting, into a raw data repository 130, raw dengue case data from a plurality of different dengue case data sources 200a-200c, the raw dengue case data being indicative of one or more dengue cases reported at one or more geographical locations for one or more time periods.

[0147] Step S2 of the method includes collecting, into the raw data repository 130, raw weather data from one or more weather data sources 200d, the raw weather data being indicative of one or more
10 weather parameters at one or more geographical locations and one or more time periods.

[0148] In step S3, a refined data repository 140 is generated based on consolidating the raw dengue case data and the raw weather data from the raw data repository 130 in terms of one or both geographical location and time. Optionally, the raw dengue case and/or weather data may be normalized to generate refined data.

15 [0149] Step S4 includes receiving, via an application programming access interface 150, API 150, an API call, for example from one or more external computing devices 300a-300c or from the dengue platform 100, requesting a dengue modelling dataset for at least one of a requested geographical location and a requested time period.

[0150] Step S5 includes providing and/or transmitting, via the API 150, the dengue modelling
20 dataset to the requester, e.g. one or more of the external computing devices 300a-300c or the dengue platform 100 itself.

[0151] The method may comprise one or more optional steps, as described in detail hereinabove. For instance, the method may comprise an optional step of computing, by an Artificial-Intelligence-based dengue outbreak prediction model 180, 302a-302c, one or more precursors associated with a
25 dengue outbreak at said at least one of the requested geographical location and the requested time period.

[0152] Figure 3 shows a block diagram illustrating operation of an exemplary dengue platform 100 and computing system 101 for predicting one or more parameters and/or precursors associated with a dengue outbreak. Unless stated otherwise, the dengue platform 100 and computing system 101 of
30 Figure 3 comprise the same features and functions as those described with reference to Figure 1. In particular, the block diagram of Figure 3 illustrates operation of the dengue platform 100 for dengue outbreak prediction and/or for training an AI-based dengue outbreak prediction model 180. In the following, reference to Figures 1, 2, and 3 is made.

[0153] Several data sources 200 for collecting raw dengue case data, raw weather data and/or
35 other data, as described hereinabove, are illustrated in Figure 3. The raw data is collected into the raw data repository 130 and converted into refined data stored in the refined data repository 140, as described above for example with reference to Figures 1 and 2.

[0154] At block 501 in Figure 3 an API call, from the dengue platform 100 itself or from one or more external computing devices 300a-300c, is received at the dengue platform's 100 API 150, which
40 provides a particular dengue modelling dataset in response. The dengue modelling dataset can then

be used to compute, by an AI-based dengue outbreak prediction model 180, 302a-302c (see Figure 1) one or more parameters and/or precursors associated with a dengue outbreak, as illustrated in block 502 of Figure 3. Accordingly, the branch from block 501 to block 502 in Figure 3 may correspond to or include steps S1 to S5 described with reference to Figure 2.

5 [0155] Alternatively or additionally, the dengue modelling dataset provided at block 501 can be further processed at block 503, for example data amputation may be applied at block 503. The amputated data may then be used at block 510 for feature engineering to define one or more predictors, based on which an AI-based dengue prediction model 180, 302a-302c can predict a dengue outbreak or parameters associated therewith.

10 [0156] At blocks 504 and 505, the data may be split into training data for training the AI-based dengue outbreak prediction model 180, 302a-302, and validation data for validating the model 180, 302a-302c. The trained and validated model 180, 302a-302c may then be provided at block 506, and further used at block 502 to compute one or more parameters and/or precursors associated with a dengue outbreak.

15 [0157] Accordingly, the dengue outbreak prediction models 180, 302a-302c can iteratively or sequentially be re-trained and/or improved by means of the dengue platform 100 or computing system 101.

[0158] While the invention has been illustrated and described in detail in the drawings and foregoing description, such illustration and description are to be considered illustrative or exemplary and not
20 restrictive; the invention is not limited to the disclosed embodiments. Other variations to the disclosed embodiments can be understood and effected by those skilled in the art and practicing the claimed invention, from a study of the drawings, the disclosure, and the claims.

[0159] As used herein, the word "comprising" does not exclude other elements or steps, and the indefinite article "a" or "an" does not exclude a plurality. The mere fact that certain measures are
25 recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage. Any reference signs in the claims should not be construed as limiting the scope.

[0160] As used herein, the phrase "being indicative of" may mean "reflecting" and/or "comprising". Accordingly, an entity/element referred to herein as "being indicative of [...]" can be synonymously or
30 interchangeably used herein with one or both said entity/element "comprising [...]" and said entity/element "reflecting [...]".

[0161] Furthermore, the terms first, second, third or (a), (b), (c) and the like in the description and in the claims are used for distinguishing between similar elements and not necessarily for describing a sequential or chronological order. It is to be understood that the terms so used are interchangeable
35 under appropriate circumstances and that the embodiments of the invention described herein are capable of operation in other sequences than described or illustrated herein.

[0162] In the context of the present invention any numerical value indicated is typically associated with an interval of accuracy that the person skilled in the art will understand to still ensure the technical effect of the feature in question. As used herein, the deviation from the indicated numerical
40 value is in the range of $\pm 10\%$, and preferably of $\pm 5\%$. The aforementioned deviation from the

indicated numerical interval of $\pm 10\%$, and preferably of $\pm 5\%$ is also indicated by the terms “about” and “approximately” used herein with respect to a numerical value.

CLAIMS

1. A computer-implemented method of providing a dengue data platform for (data-driven) predicting one or more parameters associated with a dengue outbreak, the method comprising:
- 5 collecting, in a raw data repository, raw dengue case data from a plurality of different dengue case data sources, the raw dengue case data being indicative of one or more dengue cases reported at one or more geographical locations for one or more time periods;
- collecting, in the raw data repository, raw weather data from one or more weather data sources, the raw weather data being indicative of one or more weather parameters at one or more
- 10 geographical locations and one or more time periods;
- generating a refined data repository based on consolidating the raw dengue case data and the raw weather data from the raw data repository in terms of one or both geographical location and time;
- receiving, via an application programming access interface, API, an API call requesting a
- 15 dengue modelling dataset for at least one of a requested geographical location and a requested time period; and
- providing, via the API, the dengue modelling dataset, wherein the dengue modelling dataset is processible by an Artificial-Intelligence-based dengue outbreak prediction model to predict one or more parameters associated with a dengue outbreak at said at least one of the requested
- 20 geographical location and the requested time period.
2. The method according to the preceding claim, wherein collecting the raw dengue case data into the raw data repository includes retrieving the raw dengue case data from the plurality of different dengue case data sources, and storing the raw dengue case data in a data structure
- 25 corresponding to a data structure of the respective dengue case data source.
3. The method according to the preceding claim, wherein the plurality of dengue case data sources includes at least one of one or more web pages, one or more external APIs, and one or more external databases.
- 30
4. The method according to any one of the preceding claims, wherein the raw data repository includes raw dengue case data of a plurality of heterogeneous dengue case data sources.
5. The method according to any one of the preceding claims, wherein raw dengue case data of
- 35 different dengue case data sources differ in one or more of data structure, data format, temporal granularity, time information, spatial granularity, and geographical location.
6. The method according to the preceding claim, wherein the raw data repository includes raw dengue case data indicative of dengue cases reported in one or more countries comprising Brazil,
- 40 Singapore, Malaysia, Sri Lanka, and Argentina.

7. The method according to any one of the preceding claims, wherein the raw data repository includes raw dengue case data from a plurality of data sources associated with a single country.
- 5 8. The method according to any one of the preceding claims, further comprising:
retaining, in the raw data repository, a copy of the raw dengue case data retrieved from the plurality of dengue case data sources.
9. The method according to any one of the preceding claims, further comprising:
10 re-retrieving raw dengue case data from at least one of said plurality of dengue case data sources; and
determining one or more changes in dengue cases reported for a particular geographical location and a particular time period based on inter-comparing the re-retrieved raw dengue case data and the raw dengue case data collected in the raw data repository from said at least one
15 dengue case data source.
10. The method according to any one of the preceding claims, wherein the raw dengue case data is indicative of one or more of a number of dengue cases, a number of severe dengue cases, a number of laboratory confirmed dengue cases, a number of dengue-related deaths, a country, a
20 region, a sub-region, a state, a city, a population, and a dengue serotype indicator.
11. The method according to any one of the preceding claims, wherein collecting the raw weather data into the raw data repository includes retrieving the raw weather data from at least one weather data source, and storing the raw weather data in a data structure corresponding to a data
25 structure of the respective weather data source.
12. The method according to the preceding claim, wherein the at least one weather data source includes at least one of a web page, an external APIs, and an external database.
- 30 13. The method according to any one of the preceding claims, further comprising:
retaining, in the raw data repository, a copy of the raw weather data retrieved from said one or more weather data sources.
14. The method according to any one of the preceding claims, further comprising:
35 re-retrieving raw weather data from at least one weather data source; and
determining one or more changes in one or more weather parameters reported for a particular geographical location and a particular time period based on inter-comparing the re-retrieved raw weather data and the raw weather data collected in the raw data repository from said at least one weather data source.

40

15. The method according to any one of the preceding claims, wherein generating the refined data repository includes converting the raw dengue case data into refined dengue case data based on normalizing the raw dengue case data from one or more of the plurality of dengue case data sources.

5

16. The method according to the preceding claim, wherein the raw dengue case data is normalized in terms of one or both geographical location and time.

17. The method according to the preceding claim, wherein normalizing the raw dengue case data in terms of geographical location includes determining a number of dengue cases reported for a particular geographical location with one or more predefined spatial granularities.

10

18. The method according to the preceding claim, wherein said one or more spatial granularities are associated with one or more of a country, a region in a country, and a sub-region in a country.

15

19. The method according to any one of claims 16 and 17, wherein normalizing the raw dengue case data in terms of geographical location includes determining a number of dengue cases reported for a particular geographical location with a plurality of different predefined spatial granularities.

20

20. The method according to the preceding claim, wherein normalizing the raw dengue case data in terms of time includes determining a number of dengue cases reported for a particular period of time with one or more predefined temporal granularities.

21. The method according to the preceding claim, wherein said one or more temporal granularities are associated with one or more of a day, a week, a month, a plurality of months, and a year.

25

22. The method according to any one of claims 15 to 21, wherein normalizing the raw dengue case data in terms of time includes determining a number of dengue cases reported for a particular geographical location with a plurality of different predefined temporal granularities.

30

23. The method according to any one of the preceding claims, wherein the refined data repository includes refined dengue case data and refined weather data at one or more geographical locations in one or more of a daily, weekly, a monthly and a yearly format.

35

24. The method according to any one of the preceding claims, wherein the refined data repository includes refined dengue case data indicative of one or more of a number of dengue cases, a number of severe dengue cases, a number of laboratory confirmed dengue cases, a

number of dengue-related deaths, a country, a region, a sub-region, a population, and a dengue serotype indicator.

25. The method according to any one of the preceding claims, wherein generating the refined data repository includes converting the raw weather data into refined weather data based on normalizing the raw weather data in terms of one or both geographical location and time.

26. The method according to any one of the preceding claims, wherein the raw weather data is indicative of one or more weather parameters including at least one of temperature, humidity, and precipitation, and wherein generating the refined data repository includes:

determining refined weather data based on the raw weather data, the refined weather data being indicative of one or more of a mean value, a maximum value, and a minimum value of one or more weather parameters with one or more predefined spatial granularities and/or one or more predefined temporal granularities.

15

27. The method according to the preceding claim, further comprising:
grouping and/or aggregating raw weather data according to the one or more predefined spatial granularities and/or the one or more predefined temporal granularities.

28. The method according to any one of the preceding claims, wherein generating the refined data repository includes:

deriving one or more weather-related parameters from the raw weather data, the weather-related parameters including at least one of an amount of rainfall at a geographical location within a predetermined period of time, a minimum predetermined number of consecutive days with rainfall, and a minimum predetermined number of consecutive days with a predetermined amount of rainfall per day.

29. The method according to any one of the preceding claims, wherein generating the refined data repository includes:

analyzing the raw dengue case data and/or the raw weather data in terms of missing data values; and

compensating for missing data values in the raw dengue case data and/or raw weather data based on data imputation.

30. The method according to any one of the preceding claims, wherein generating the refined data repository includes:

analyzing the raw dengue case data and/or the raw weather data in terms of outliers; and removing one or more outliers from the raw dengue case data and/or the raw weather data.

31. The method according to any one of the preceding claims, wherein the received API call includes a request from an external computing device to obtain a dengue modelling dataset for a requested geographical location at a requested spatial granularity, and wherein the method further comprises:

5 retrieving refined dengue case data and/or refined weather data from the refined data repository for the requested geographical location and with the requested spatial granularity; and
 providing the dengue modelling dataset based on the refined dengue case data and/or the refined weather data.

10 32. The method according to any one of the preceding claims, wherein the received API call includes a request from an external computing device to obtain a dengue modelling dataset for a particular country, region or sub-region, and wherein the method further comprises:

 retrieving refined dengue case data and/or refined weather data from the refined data repository for the requested country, region or sub-region; and
15 providing the dengue modelling dataset based on the refined dengue case data and/or the refined weather data.

33. The method according to any one of the preceding claims, wherein the received API call includes a request from an external computing device to obtain a dengue modelling dataset for a requested time period at a requested temporal granularity, and wherein the method further comprises:

20 retrieving refined dengue case data and/or refined weather data from the refined data repository for the requested time period and with the requested temporal granularity; and
 providing the dengue modelling dataset based on the refined dengue case data and/or the refined weather data retrieved from the refined data repository.
25

34. The method according to any one of the preceding claims, further comprising:
 converting the received API call into a database query for querying the requested dengue modelling dataset from the refined data repository.

30 35. The method according to the preceding claim, wherein the database query is a location-based database query, a time-based database query, or a spatio-temporal database query.

36. The method according to any one of the preceding claims, wherein the dengue modelling dataset is indicative of a number of dengue cases at a requested geographical location and/or for a requested period of time.
35

37. The method according to any one of the preceding claims, wherein the dengue modelling dataset is indicative of one or more of a number of severe dengue cases, a number of laboratory confirmed dengue cases, a number of dengue-related deaths, a population, a dengue case data
40

source, a dengue serotype indicator, a data quality, a country, a region, a sub-region, a state, a temporal granularity, and a spatial granularity.

38. The method according to any one of the preceding claims, wherein the dengue modelling
5 dataset is indicative of one or more of a minimum temperature, a maximum temperature, an average
temperate, a humidity, and a precipitation at the requested geographical location and/or for a
requested period of time.

39. The method according to any one of the preceding claims, wherein the dengue modelling
10 dataset is indicative of one or more of a country of a weather station, a measurement location, a
county of a weather station, a municipality of a weather station, a city of a weather station, a
postcode of a weather station, a latitude of a weather station, and a longitude of a weather station.

40. The method according to any one of the preceding claims, wherein the dengue modelling
15 dataset is indicative of one or more of an amount of rainfall, a minimum predetermined number of
consecutive days with rainfall, and a minimum predetermined number of consecutive days with a
predetermined amount of rainfall per day.

41. The method according to any one of the preceding claims, further comprising:
20 training the Artificial-Intelligence-based dengue outbreak prediction model using the dengue
modelling dataset.

42. The method according to any one of the preceding claims, further comprising:
providing the dengue modelling dataset as input to an AI-based dengue outbreak prediction
25 model to predict one or more parameters associated with a dengue outbreak.

43. The method according to any of the preceding claims, further comprising:
predicting, based on processing the dengue modelling dataset with an AI-based dengue
outbreak prediction model one or more parameters associated with a dengue outbreak.

30 44. The method according to the preceding claim, further comprising:
computing one or more of a probability of an onset of a dengue outbreak, a time of an onset
of a dengue outbreak, a geographical location of a dengue outbreak, an epicenter of a dengue
outbreak, a number of expected cases of a dengue outbreak, a number of expected cases of a
35 dengue outbreak within a predetermined period of time, and a severity measure for a dengue
outbreak.

45. The method according to any one of the preceding claims, further comprising collecting
further dengue-related and/or weather-related data from one or more further data sources including

social media platforms, world bank data, open data, research data, scientific publications, publication data, social data, building density data, vector-related data, mosquito data.

46. A computer program, which, when executed by one or more processors of a computing system, instructs the computing system to perform steps of the method according to any one of the preceding claims.

47. A non-transitory computer-readable medium storing a computer program according to the preceding claim.

48. A computing system for predicting one or more parameters associated with a dengue outbreak, wherein the computing system includes processing circuitry with one or more processors and is configured to perform steps of the method according to any one of claims 1 to 45.

49. The computing system according to the preceding claim, comprising:
a raw data repository storing the raw dengue case data from a plurality of different dengue case data sources, and storing raw weather data from one or more weather data sources; and
an API configured to receive an API call from an external computing device and configured to provide a dengue modelling dataset in response to said API call.

50. The computing system according to any one of claims 48 and 49, further comprising an Artificial-Intelligence-based dengue outbreak prediction model configured to receive the dengue modelling dataset as input and to compute one or more parameters associated with a dengue outbreak based on processing the dengue modelling dataset.

51. A computer-implemented method executed by data processing hardware that causes the data processing hardware to perform operations comprising:

collecting, at a raw data repository, raw dengue case data from a plurality of different dengue case data sources, the raw dengue case data indicative of one or more dengue cases reported at one or more geographical locations for one or more time periods;

collecting, at the raw data repository, raw weather data from one or more weather data sources, the raw weather data indicative of one or more weather parameters at the one or more geographical locations and the one or more time periods;

generating a refined data repository based on the raw dengue case data and the raw weather data;

receiving, via an application programming access interface (API), an API call requesting a prediction of a parameter associated with a dengue outbreak, the API call comprising at least one of a requested geographical location or a requested time period;

generating a dengue modelling dataset from the refined data repository based on the at least one of the requested geographical location or the requested time period;

training a dengue outbreak prediction model using the dengue modelling dataset;
generating the prediction using the trained dengue prediction model; and
providing, via the API, the prediction.