

(51) International Patent Classification:
G16H 10/40 (2018.01)(21) International Application Number:
PCT/EP2021/051997(22) International Filing Date:
28 January 2021 (28.01.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

20154563.9	30 January 2020 (30.01.2020)	EP
20154565.4	30 January 2020 (30.01.2020)	EP
20172153.7	29 April 2020 (29.04.2020)	EP
20182720.1	26 June 2020 (26.06.2020)	EP
20182885.2	29 June 2020 (29.06.2020)	EP
20197902.8	23 September 2020 (23.09.2020)	EP

(71) Applicant: **MEDICUS AI GMBH** [AT/AT]; Wehleweg
9 / 53, 1030 Wien (AT).

(72) Inventors: **AL HAKIM, Baher**; c/o Medicus AI GmbH, Wehleweg 9 / 53, 1030 Wien (AT). **KAWAS, Mouhamad**; c/o Medicus AI GmbH, Wehleweg 9 / 53, 1030 Wien (AT). **ALKHATIB, Bassel**; c/o Medicus AI GmbH, Wehleweg 9 / 53, 1030 Wien (AT). **NEHME, Nadine**; c/o Medicus AI GmbH, Wehleweg 9 / 53, 1030 Wien (AT).

(74) Agent: **STELLBRINK & PARTNER PATENTAN-
WÄLTE MBB**; Widenmayerstr. 10, 80538 Munich (DE).

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

(54) Title: MEASUREMENT DATA PROCESSING

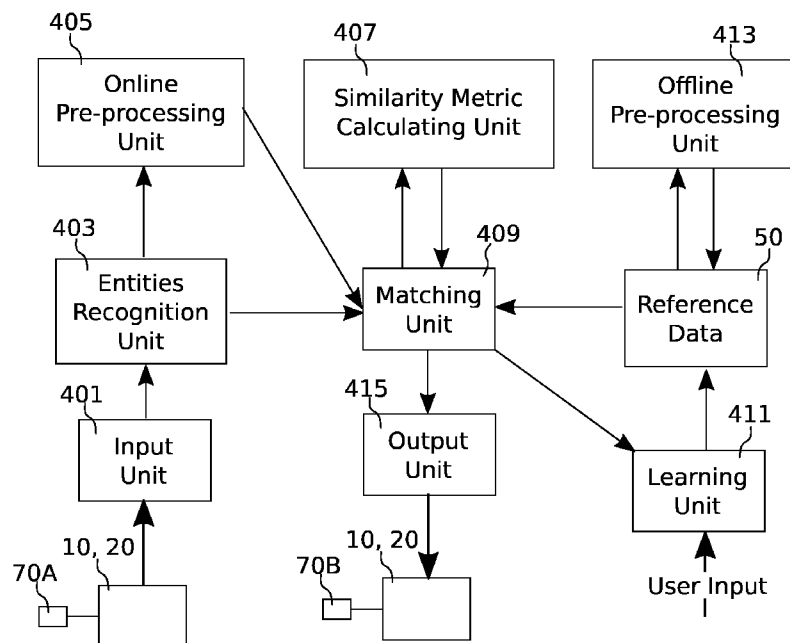


Fig. 10

(57) Abstract: The present invention relates to a data processing system comprising an input unit configured to obtain measurement related data configured according to a first code and a matching unit (409) configured to compare the obtained measurement related data (10, 20) with reference data (50) and to configure the measurement related data (10, 20) according to a second code (70B) based on the comparison. In addition, the present invention relates to a method that can be carried out by the data processing system and to a communication system that can comprise the data processing system.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

Measurement data processing

The present invention relates to the field of processing measurement data, particularly measurement data generated by bio-medical laboratories.

In modern measurement technology, measurement data are typically stored and processed digitally. A measure, the result of a quantitative measurement process, is typically stored as a number together with a unit corresponding to the dimension of the result and a specifier indicating the type of measurement. For example, a measurement of a length can be stored as 59 mm, 5,9 cm or 2,323 inches. However, for many measurements or measurement processes, further specifications are necessary. For example, measuring a height or weight of a car, it is necessary to specify under which circumstances the measurement is performed. For example, air pressure in the tires, load and even modifications of the vehicle may have an impact on the measurements. The same applies to other measures, e.g., a measure of a viscosity of a fluid is typically only meaningful together with an indication of the temperature of the fluid.

In medicine and biology, the same principles apply. For example, a concentration of a substance can be measured in different tissues or body fluids (i.e. in different components or samples). Hence, even if the concentration is always indicated in the same dimension, i.e. unit, e.g., mass per volume, this value is not meaningful without an indication of the body fluid or tissue (i.e. component) in which it was measured. Furthermore, and in contrast to most measuring methods in mechanical or electrical engineering, the measuring method in medicine and biology may impact a result. For example, when determining a concentration of a substance, such as cholesterol high-density lipoprotein, the measuring method may lead to different measurement values for a same sample or samples from a same user.

Also, a reference or name of a generated measurement may be relevant, so as to allow the identification of the measurement and processing of the values by means of a data-processing system.

In other words, typically a measurement is stored, processed or communicated with, e.g., a name of the measurement, a number, a value, a range, a range specifier, a unit, a component, a sample on which the measurement is performed and a used method. For example, in medical or biological laboratories or hospitals, the corresponding laboratory or hospital information systems may store, process and/or communicate measurements as discussed above.

Moreover, generally, different nodes (e.g. measurement systems, laboratories, clinics, practices, hospitals, sensor devices, measurement devices, laboratory or hospital information systems) may represent measurements differently or using different standards. In many instances, nodes use their own proprietary standards to store, process or communicate measurements. For example, they may use different names and units for the same measures and may generate measures based on different standards.

However, this may lead to limitations and disadvantages.

As an initial matter, when processing a plurality of measurements, it can be required and essential to have the measurement values (and the other data corresponding to the measurement) in a comparable form. This is a problem well-known in multiple fields of science and technology, such as physics, chemistry, engineering, data science, quality control in production, or analysis of lab results in medicine, wherein it is typically required for the data to be homogenized before processing.

In addition, most measurement systems that comprise a data-processing system accept only a limited number of types of instructions as regards measurements for one sample. Some even only accept one type of instructions.

These limitations can lead to disadvantages regarding the interoperability of different measurement performing/requesting systems and analysis systems, e.g., different laboratories and receivers of the measured values. Due to such limitations, e.g., a quality control system or a physician typically only cooperates with one or few laboratories. Results from other laboratories or measurement systems may be unusable, as they may not be comparable to results from prior measurements, defined limits for an "acceptable"-range, and so on.

Also, especially when it comes to highly specialised measurements or mass-processing of samples that have to be analysed, i.e. for which measurements have to be made, in the past decade, central processing of samples in central laboratories, which are sometimes even specialised to single measurements or a field of measurements, has become more common. Hence, clients send their samples to a service provider, who may then redistribute these samples to an appropriate facility for performing the requested measurement. This can obviously include multiple parties on the client side, the (intermediary) service provider side and the laboratory side. They may each have their own data-processing systems, and these data-processing systems may at least partially work with different standards.

EP1200839A1 discloses an integrated clinical laboratory software system for testing a specimen.

US 2011/0119309 A1 discloses a gateway enabling medical (including genetic and genomic) laboratories and health care providers (collectively "clients") to communicate electronic messages with each other without developing and maintaining an interface for each peer.

WO2003040697A1 discloses a method and devices for the cross-referencing of identification of object supports, for microtomed analytical samples still to be mounted thereon, with identification information for a support of a tissue sample which is not yet microtomed. The conventional problem of cross-referencing is improved in a simple manner, whereby the identification information for the support is automatically generated during the very allocation in the microtome and an identification, corresponding thereto, is automatically transferred to at least one object support and that finally said object support, provided with the identification is given for the application of the microtomed tissue sample at the moment when a microtomed tissue sample must be applied to an object support.

As can be seen, an interoperability of different measurement systems and systems receiving or processing their measurements is of increasing importance. Also, corrections or other interactions of the measurement results may be necessary in such processes and may disadvantageously introduce errors, which are to be avoided or minimized. It is also obvious to the person skilled in the art, that the above-mentioned constellation may be disadvantageous as regards a flexibility of the overall system, e.g. when a participant of the system changes their inventory from one (established) version to another (established) version, since there are several dependencies to the systems of other participants.

It is therefore an object of the invention to overcome or at least alleviate the shortcomings and disadvantages of the prior art. More particularly, it is an object of the present invention to provide a system, method and computer program product for processing measurement related data.

It is an optional object of the present invention to provide a system, method and computer program product for providing measurement result data in an encoding that can be processed by another data-processing system.

It is an optional object of the present invention to provide a system, method and computer program product for providing measurement result data in a standard or universal encoding.

It is another optional object of the present invention to provide a system, method and computer program product for providing instructions for a measurement system in an encoding that the measurement system can process.

It is another optional object of the present invention to increase an interoperability between different measurement systems and data-processing systems.

In a first aspect the present invention relates to a method of processing measurement related data. The method comprises a data processing system obtaining measurement related data configured according to a first code. The method further comprises the data processing system comparing the obtained measurement related data with reference data. Moreover, the method comprises the data processing system configuring the measurement related data according to a second code, based on the comparison.

The measurement related data can for example be a laboratory report, such as, a medical laboratory report.

A code, as used herein, i.e., the first code and the second code, can be a system of rules or instructions for configuring the measurement related data. For example, the code can comprise a system of measurement identifiers (e.g. a system of names or terms) that can be used to refer to measurements. In other words, the code can comprise a respective measurements nomenclature. In addition, the code can define a structure for configuring the measurement related data. For example, the code may define a tabular structure which can be filled with measurement data. In general, the code may define a data structure, the values of which can correspond to the measurement data. Further still, the code may define an order of the measurement related data. The code may also define a format of the measurement related data. For example, the code may comprise instructions (e.g. computer instructions) for configuring the measurement related data as image data, text file and/or according to a file format.

The second code can also be referred to as the target code.

That is, the present method comprises utilizing a data processing system to obtain measurement related data configured according to a first code, compare them with reference data and based thereon configure the measurement related data according to a second code.

The present method can increase the interoperability between two or more communicating nodes in a communication network. This can be particularly the case when nodes in a communication network utilize different codes to generate and/or encode data. For example, a first node may utilize the first code and a second node may utilize the second code, wherein the first code can be different from the second code. While otherwise the communication between the nodes may not be feasible, the present method can facilitate the communication between the two nodes.

For example, a physician may request a measurement to be performed on a sample from a patient. The physician may use an internal code to configure the measurement request. For example, the physician may refer to the measurement with a non-standard name or with a name internally used by the physician and respective staff. A laboratory device receiving the measurement request from a physician's computing device (e.g. from a computer used by the physician) may utilize another code to refer to measurements. For example, the hospital information system (HIS) used by the doctor can utilize the first code to configure measurement related data and the laboratory information system (LIS) of the laboratory requested to perform the measurement can utilize the second code. As such, the data processing system can facilitate the communication between the HIS and the LIS.

Moreover, the present invention can be advantageous as it can provide a data processing system for configuring the measurement related data from the first code to the second code. Generally, the measurement related data can be communicated from a sending node to a second node. For example, from a physician's computer to a laboratory device. It is typically the case, that the nodes can have limited computational power. That is, generally the nodes can be localized in laboratories or medical practitioner's offices. Due to space and cost limitations, these nodes can have limited computational power. As such, they may not be able e.g., to obtain measurement related data configured according to a first code and configured them according to a second code, or to a code that the node utilizes. However, this can be efficiently performed by a data processing system configured to carry out the method of the present invention. The data processing system can comprise sufficient computational units and/or can be customized to perform the method. Hence, the method can be efficiently executed and the communication between the nodes facilitated. It will be understood, that the data processing system may serve more than one pair of communication nodes in a communication system.

In addition, the present invention can generally increase the accuracy of performing measurements. Typically, due to different codes (e.g. nomenclatures) used by different nodes, there can be ambiguity on the identifiers that different nodes may use to refer to measurements. As such, while a first node may request a second node to perform a measurement, a second node may perform another measurement or may perform the measurement not as requested, e.g., using a different method. Again, the present invention can alleviate such issues.

Further still, the present method can facilitate configuring measurement related data according to standard codes. The present invention achieves this by selecting the second code to be a standard code. For example, laboratory reports can be configured according to the LOINC (Logical Observation Identifiers Names and Codes) universal standard. This can facilitate the use of the measurement related data (e.g. laboratory report) in multiple applications.

Alternatively or additionally, the present method can facilitate configuring the measurement related data according to a code used by a receiving node. For example, the measurement related data can be communicated from a sending node to a receiving node. Moreover, the measurement related data can be generated by the sending node using a code corresponding to or being internally used by the sending node. The present method can comprise utilizing the data processing system to configure the measurement related data according to a different code which can correspond to or can be internally used by the receiving node.

In other words, the present method can be advantageous as it can allow configuring measurement related data according to at least one target code (i.e. second code) which can be a standard code and/or a non-standard code used by a receiving node.

The obtained measurement related data can comprise at least one measurement identifier which can be configured to identify a measurement. For example, the measurement identifier can be a name of a measurement (e.g. a biomarker name). The measurement identifier can comprise at least one of symbols, unique identification sequences, names, short names, long names and abbreviations.

In such embodiments, the method can comprise detecting at least one measurement identifier in the obtained measurement related data. That is, the measurement related data can generally be unstructured data. In other words, it may not be readily possible, e.g., by the

data processing device, to determine which parts of the measurement related data are measurement identifiers. Detecting the at least one measurement identifier can facilitate structuring the measurement related data. For example, parts of the measurement related data that correspond to measurement identifiers can be labeled as such. In addition to facilitating the structuring of the measurement related data, detecting at least one measurement identifier can further facilitate comparing the measurement related data with reference. For example, the detected measurement identifiers in the obtained measurement related data can be compared with identifier references comprised in the reference data to determine matching/replacing matching identifier references, which can be used to configure the measurement related data according to the second code.

Preferably, the method comprises the data processing system detecting the at least one measurement identifier in the obtained measurement related data. That is, the step of detecting at least one measurement identifier in the obtained measurement related data can be carried out by the data processing system.

Preferably, the method can comprise detecting each of the at least one measurement identifiers comprised in the obtained measurement related data.

The obtained measurement related data can comprise a plurality of data portions.

In some embodiments, the obtained measurement related data can consist of a plurality of data portions.

The data portions can be non-intersecting portions of the obtained measurement related data.

For example, the data portions can be blocks of data within the measurement related data. That is, the measurement related data can be arranged in blocks of data or data block.

In such embodiments, the method can comprise detecting at least one data portion in the obtained measurement related data. Preferably, the method can comprise detecting each of the at least one data portion comprised in the obtained measurement related data. Detecting the at least one data portion can facilitate structuring or determining the structure of the measurement related data. Moreover, in embodiments of the present method, detecting the at least one data portion can facilitate detecting at least one measurement identifier.

The data portions can correspond to respective lines of the obtained measurement related data.

In some embodiments, the obtained measurement related data can comprise portion delimiters configured to specify the boundaries of the data portions. Thus, the portion delimiters can facilitate detecting the at least one data portion.

That is, the data portions can be separated from each other using portion delimiters.

The portion delimiters can be new line characters. This can be the case if the data portions correspond to respective lines of the obtained measurement related data. However, it will be understood that is exemplary.

In some embodiments, detecting at least one measurement identifier can comprise detecting at least one data portion of the obtained measurement related data that comprises a measurement identifier.

Preferably, to detect at least one data portion of the obtained measurement related data that comprises a measurement identifier, the method can comprise detecting each data portion of the obtained measurement related data.

Moreover, detecting at least one data portion of the obtained measurement related data that comprises a measurement identifier can comprise determining for each of the detected data portions whether it comprises a measurement identifier.

That is, in some embodiments, the method can detect measurement identifiers in the measurement related data by detecting data portions in the measurement related data and by classifying the detected data portions as comprising or not comprising measurement identifiers. The method can further comprise utilizing the data portions classified as comprising measurement identifiers to detect the at least one measurement identifier. Thus, instead of searching directly for measurement identifiers over the entire measurement related data, in some embodiments, firstly, data portions containing at least one measurement identifier can be detected and based thereon the measurement identifiers can be detected. This can reduce the number of computations and/or the time required to detect a measurement identifier.

The measurement related data can comprise a plurality of data elements.

In embodiments wherein the measurement related data can comprise at least one data portion, each data portion can comprise at least one data element. That is, the measurement related data can be arranged in data portions, each comprising data elements. For example, the data portions can be respective lines and the data elements can be words.

The method can further comprise detecting the plurality of data elements. This can facilitate structuring and/or determining the structure of the measurement related data, which in turn facilitates the processing of the measurement related data, e.g., by the data processing system.

Each data element can comprise a plurality of data bits, at least one byte of data and/or at least one character, such as, at least one ASCII character.

In some embodiments, the method can comprise associating to each data element a plurality of data bits, at least one byte of data and/or at least one character, such as, at least one ASCII character.

Moreover, each of the at least one measurement identifiers can comprise at least one data element.

In some embodiments, detecting at least one measurement identifier can comprise determining for each data element in the measurement related data whether it corresponds to a measurement identifier. This step of the method can also be referred to as direct detection. That is, the measurement identifiers in the measurement related data can be searched for directly.

In such embodiments, determining for each data element in the measurement related data whether it corresponds to a measurement identifier can comprise comparing each data element and/or a sequence of data elements with a plurality of measurement identifier references. That is, in some embodiments plurality of measurement identifier references can be provided. For example, the plurality of measurement identifier references can comprise a plurality of measurement identifiers expected to be comprised in measurement related data. Thus, comparing each data element and/or a sequence of data elements with each (or at least some) of the plurality of measurement identifier references can allow the detection of the measurement identifiers.

The reference data can comprise the plurality of measurement identifier references. That is, the method can comprise providing reference data comprising a plurality of measurement identifier references. In other words, the reference data can comprise a plurality of measurement identifiers, expected to be comprised in the measurement related data. On the one hand, this can facilitate the detection of measurement identifier references, e.g., using the direct detection step as discussed above. On the other hand, this can facilitate configuring the measurement related data according to the second code, as will be discussed further below.

The data elements can be identifiable.

For example, the measurement related data can comprise element delimiters configured to specify boundaries of the data elements.

In some embodiments, the method can comprise detecting at least one measurement indicator in the obtained measurement related data. As discussed, typically a measurement can be provided with an identifier and with data characterizing the measurement. The latter can be, e.g., numbers, values, units, ranges and range specifiers. Detecting the at least one measurement indicator can be advantageous as it can facilitate detecting at least one measurement identifier, e.g., through an indirect detection step, as will be discussed further below. In addition, detecting the at least one measurement indicator can facilitate structuring and/or determining the structure of the measurement related data. That is, it can facilitate determining parts of the measurement related data that can further characterize a measurement, such as, measurement results, units, ranges and range specifiers.

The at least one measurement indicator can comprise at least one of numerical values, ordinal values, nominal values, qualitative data, quantitative data, unit of measurement, range and range specifier. In other words, detecting at least one measurement indicator can comprise searching for the presence of numbers, values (e.g. numerical or text values), units, ranges and/or range specifiers. This can be performed fast and may require little computation to be performed. For example, to detect at least one measurement indicator it can be sufficient to detect one of the numbers from 0 to 9, one of the units from a set of possible units and/or one of the range specifiers from a set of possible range templates. The range templates can, for example, comprise at least one regular expression corresponding to and/or configured to

facilitate detecting ranges. The set of possible units and the set of possible range templates can be comprised by the reference data, as discussed further below.

Detecting at least one measurement identifier can comprise detecting at least one measurement identifier based on the detection of at least one measurement indicator. This step of the method can be referred to as indirect detection, as the measurement identifiers can indirectly be detected based on the detection of at least one measurement indicator. The indirect detection can be a particularly efficient way of detecting measurement identifiers.

As discussed, the measurement identifiers can be directly detected by searching for the presence of at least one of a plurality of measurement identifier references. However, as the plurality of measurement identifier references can be very large, e.g., it can comprise all the possible identifier references, detecting measurement identifiers this way can be inefficient, time consuming and can require a large number of computations. On the other hand, detecting at least one measurement indicator, can require less computations. For example, the size of all possible units and all possible range templates that can be used to detect measurement indicators can be smaller than the set of all measurement identifiers. As such, the indirect detection step can generally be more time and resource efficient than the direct detection method.

In some embodiments, the method can comprise determining the location of the at least one measurement indicator in the obtained measurement related data. This can comprise at least one of determining an index in a list, indices in a multi-dimensional vector, position on an image, position on a table. In some embodiments, determining the location of the at least one measurement indicator can comprise determining on which of the data portions, a measurement indicator is present.

Detecting at least one measurement identifier can comprise detecting at least one measurement identifier based on the location of the at least one measurement indicator in the obtained measurement related data. That is, the method can comprise utilizing the location of at least one measurement indicator in the measurement related data to detect at least one measurement identifier.

For example, data that precedes or follows the at least one measurement indicator can be determined (or hypothesized with a respective certainty measure, e.g., likelihood) to be a measurement identifier.

In some embodiments, the method can comprise determining whether the at least one data portion comprises at least one measurement indicator. This can for example be performed based on the location of the at least one measurement indicator.

Detecting at least one data portion of the obtained measurement related data that comprises a measurement identifier can comprise determining that at least one data portion comprises a measurement identifier if the at least one data portion comprises at least one measurement indicator. In other words, if it can be determined that a data portion comprises a measurement indicator then it can further be determined that the data portion comprises a measurement identifier. This can be based on the rationale that the measurement indicator and the measurement identifier can generally be positioned close to each other, e.g., in the same data portion. As such, the detection of the measurement identifiers in the measurement related data can be facilitated by determining parts or portions of the measurement related data that can be associated with a high likelihood of comprising a measurement identifier. This can, for example, allow searching for the measurement identifiers only on data portions comprising at least one measurement indicator, instead of searching the entire measurement related data. Thus, time and/or computational resources can be saved.

In some embodiments, the method can comprise upon detecting at least one measurement indicator in a data portion, determining that the remaining data in the data portion comprise the measurement identifier. As discussed, a data portion can comprise at least one data element. Moreover, for one or some data portions it can be determined that they comprise at least one measurement indicator. That is, it can be determined that at least one of the data elements comprised by a data portion can be measurement indicator(s). In such instances, it can be determined that the rest of the data elements in the data portion can be a measurement identifier. For example, it can be determined that the data elements preceding the measurement indicator can be a measurement identifier.

For example, a line in a laboratory report can comprise a biomarker name followed by a number, a value (which can be a numerical value and/or a text value), a unit, a range and a range specifier. As discussed, the numbers, values, units and ranges can be detected in the laboratory report. As such, the respective line wherein the numbers, units and/or ranges are detected can be classified as biomarkers lines (i.e. as data portions comprising at least one measurement identifier). The rest of the data elements (e.g. words) comprised in the biomarker lines – except from the detected numbers, units and ranges (which are

measurement indicators) – can be determined with a high likelihood to be a biomarker (i.e. a measurement identifier).

In some embodiments, the method comprises upon detecting at least one measurement indicator in a data portion, determining that the remaining data in the data portion is the measurement identifier. That is, in some embodiments, not only it can be determined that the remaining data in a data portion wherein a measurement indicator is detected can comprise a measurement identifier, but it can be specified that the remaining data is the measurement identifier.

As will be discussed further, in embodiments of the present invention, the determination that the remaining data in a data portion wherein a measurement indicator is detected can comprise or can be the measurement identifier can be further validated, e.g., during the matching step discussed below. In other words, the said determination can be a hypothesis and, in some embodiments, it can be associated with a respective likelihood of being the true hypothesis. For example, the said likelihood can be calculated based on the number and type of the measurement indicators detected in data portion. For example, the mere detection of a number in a data portion may yield a hypothesis that a measurement identifier can be present in the data portion with a lower likelihood than the detection of a number and a unit and with an even lower likelihood than the detection of a number, unit, and range specifier.

The at least one measurement indicator can comprise a characteristic type. As discussed, the measurement indicator can be configured to characterize a measurement. For example, the measurement indicator can be configured to indicate the result of the measurement in a qualitative or quantitative form, the unit of the measurement and a range indicating the possible values of the measurement result. As such, the measurement indicator can be of different types. It will be understood that the above is not an exhaustive list of the measurement indicators.

The characteristic type of the at least one measurement indicator can comprise at least one of numerical value, ordinal value, nominal value, qualitative data and quantitative data. These can typically be used to indicate the value or result of a measurement. Alternatively or additionally, the characteristic type of at least one measurement identifier can comprise a unit of measurement, and a range specifier. Again, it will be understood that the above is not an exhaustive list of the characteristic types corresponding to the measurement indicators.

In some embodiments, detecting at least one measurement indicator in the obtained measurement related data can comprise determining the characteristic type of the at least one measurement indicator. This can, for example, intrinsically be determined during the detection of the at least one measurement indicator. For example, if a data element is determined to be a unit, then the data element can be determined to be a measurement indicator with a unit as a characteristic type. In other words, in some embodiments, the detection of a measurement indicator and the characteristic type corresponding to the measurement indicator can be performed simultaneously.

Detecting the type of the measurement indicator can be advantageous as it can facilitate the detection of data portions comprising a measurement identifier, particularly, when regular expressions are used to detect data portions comprising a measurement identifier. In addition, it can facilitate determining a likelihood that the data portion comprises a measurement identifier. As discussed, the mere detection of a number in a data portion may yield a hypothesis that a measurement identifier can be present in the data portion with a lower likelihood than the detection of a number and a unit and with an even lower likelihood than the detection of a number, unit, range and range specifier.

Detecting at least one measurement indicator in the obtained measurement related data can comprise using regular expressions. Alternatively or additionally, the regular expressions can be used to detect data portions that can comprise at least one measurement identifier (e.g. to detect biomarker lines in a laboratory report). Generally, the data portions can be compared with predetermined regular expressions. This can facilitate determining whether the data portion comprises a measurement identifier (e.g. whether a line in a laboratory report is a biomarker line) and/or whether the data portion comprises a measurement indicator. In addition, by matching a data portion with a regular expression, it can be determined which part of the data portion correspond to a measurement identifier and which part(s) correspond to measurement indicator(s).

As discussed, a measurement can generally be represented with a measurement identifier (which generally is compulsory) and further optional data (i.e. measurement indicators), which can be numerical value, ordinal value, nominal value, qualitative data, quantitative data, unit of measurement, range and range specifier. However, depending on the measurement related data and on the code according to which the measurement related data can be configured, the order and the amount of data used to identify and characterize a measurement can be different. For example, a first exemplary measurement related data can comprise for each

measurement only a measurement identifier. A second exemplary measurement related data can comprise for each measurement a measurement identifier and a unit. A third exemplary measurement related data can comprise for each measurement a measurement identifier, a value and a unit. A fourth exemplary measurement related data can comprise for each measurement a measurement identifier, a unit and a value (in this order). The present technology can utilize predetermined regular expression to consider different arrangements of the measurement identifiers and measurement indicators.

In some embodiments, the present technology can utilize a large number of regular expression (e.g. at least 100 regular expressions), each corresponding to a respective arrangement of the measurement identifiers and measurement indicators. Preferably, the regular expressions can cover all possible arrangements of the measurement identifiers and measurement indicators in the measurement related data. This can be particularly advantageous as it can facilitate the detection of the measurement identifiers and measurement indicators in measurement related data configured according to a large variety of codes. That is, the use of regular expressions and more particularly the use of regular expressions covering a plurality (preferably all) of possible arrangements of the measurement identifiers and measurement indicators can facilitate detecting the measurement identifiers and measurement indicators in measurement related data configured according to an arbitrary code.

For example, a laboratory report can comprise the line "High Density Lipoprotein Cholesterol 100 mg/L 140-200". This line can be matched with the regular expression "BiomarkerName NumericalValue Unit From-To-Range-Min-Max".

In some embodiments, heuristic rules can be utilized to further facilitate the detection of measurement indicators and/or the detection of data portions comprising at least one measurement identifier. The heuristic rules can be utilized in addition to the regular expression. More particularly, the heuristic rules can be used to solve ambiguities or conflicts that can occur while matching data portions with the regular expressions. Some heuristic can, for example, be used when a data portion matches more than one regular expression (in this case the longest match can be considered), when spaces are used with numbers (in this case x space y can be a range), when a unit contains range-like pattern, e.g., 10~9/L (in this case x~y can be a range), when a data element can be biomarker name (i.e. measurement identifier) or a unit (i.e. measurement indicator) and when a biomarker name contains text value, e.g., high, red. It will be understood, that the above ambiguities are provided for exemplary reasons only.

In some embodiments, detecting at least one measurement indicator can comprise using a string-searching algorithm.

Each of the at least one measurement indicators can comprise at least one data element and wherein detecting at least one measurement indicator can comprise determining for each data element in the measurement related data whether it corresponds to a measurement indicator.

Determining for each data element in the measurement related data whether it corresponds to a measurement indicator can comprise comparing each data element and/or a sequence of data elements with a plurality of reference characteristics.

The reference data can comprise the plurality of reference characteristics. For example, the reference data can comprise for each measurement a plurality of possible units and/or range templates that can be used to indicate or further characterize the measurement.

In some embodiments, the method can comprise a pre-processing step.

The pre-processing step can comprise pre-processing the at least one detected measurement identifier.

The pre-processing step can comprise pre-processing the reference data.

This can facilitate the step of the data processing system comparing the obtained measurement related data with reference data. The pre-processing step can be advantageous as it can increase the homogeneity between the measurement related data and the reference data. This can be achieved by performing the pre-processing step (e.g. once offline) to the reference data and storing the reference data after pre-processing step. In addition, the pre-processing step can be performed (online) to the measurement related data. The pre-processing step can be particularly advantageous to reduce the false-negative rate while matching the measurement identifiers detected in the measurement related data with the reference data.

The pre-processing step can be performed before the step of the data processing system comparing the obtained measurement related data with reference data.

The data processing system can perform the pre-processing step. That is, the pre-processing step can be a computer implemented step (i.e. can comprise computer instructions) which can be executed by the data processing system.

The pre-processing step can comprise performing a data cleaning of the at least one detected measurement identifier.

The data cleaning step can comprise detecting at least one data element, comprised by the at least one detected measurement identifier, which does not facilitate the identification of a measurement. That is, the data cleaning step can comprise removing from the measurement identifiers, data elements with a low relevance on identifying a measurement.

The data elements which do not facilitate the identification of a measurement may also be referred to as irrelevant data elements.

Detecting at least one data element which does not facilitate the identification of a measurement may comprise providing a data cleaning database and determining for each data element comprised by the at least one detected measurement identifier whether it is part of the data cleaning database. The data cleaning database can comprise a plurality of possible irrelevant data elements (e.g. a plurality of stop-words).

The reference data can comprise the data cleaning database.

The data cleaning database can comprise a plurality of stop-words, symbols and punctuation marks.

The method can comprise the data processing system comparing the obtained measurement related data with reference data without utilizing the at least one data element which does not facilitate the identification of a measurement. This can be advantageous as the irrelevant data elements may act as artefacts, thus increasing the likelihood of yielding false results during the comparison.

The pre-processing step can comprise removing from the measurement identifiers the at least one data element which does not facilitate the identification of a measurement.

In some embodiments, pre-processing the at least one detected measurement identifier can comprise counting the number of data elements comprised by the at least one detected measurement identifier.

Thus, a data element count can be determined for some or each of the at least one detected measurement identifier.

Similarly, a data element count can be determined for each identifier references that can be comprised by the reference data.

The data element count can facilitate comparing the measurement related data with the reference data. As it will be discussed, heuristic rules based on the data element counts can be utilized to reduce the number of comparisons between the measurement related data and the reference data.

Counting the number of data elements comprised by the at least one detected measurement identifier can comprise skipping the detected data elements which do not facilitate the identification of a measurement. In other words, the irrelevant data elements are not counted.

For example, counting the number of data elements comprised by the at least one detected measurement identifier can be performed after removing the data elements which do not facilitate the identification of a measurement from the measurement identifiers.

The pre-processing step can comprise replacing at least one data element comprised by the at least one detected measurement identifier with an equivalent data element.

Replacing at least one data element comprised by the at least one detected measurement identifier with an equivalent data element can comprise utilizing an equivalent data elements database and wherein the equivalent data elements database comprises a plurality of data elements each associated with at least one equivalent data element.

For example, the equivalent data elements database can comprise a synonyms dictionary.

Utilizing an equivalent data elements database can comprise searching the equivalent data elements database for at least one data element comprised by the at least one detected measurement identifier.

The data element and the equivalent data element can convey same or similar information.

It will be understood that each data element of the measurement related data can be configured to convey information.

The data element and the equivalent data element can be synonyms.

The data element and the equivalent data element can comprise different grammatical forms of the same word.

The equivalent data element can comprise a word stem (i.e. root) of the data element.

A stem or root of a word can be a form of a word before any inflectional affixes are added.

The pre-processing step can comprise performing word stemming on the data elements.

Generally, information conveyed by a data element can generally be comprised in a part of a data element. For example, the data element can be a word, and the root or stem of the word can typically comprise most of the information conveyed by the word. On the other hand, two data elements can convey the same information however due to affixes they can differ. Thus, by stemming the data elements, i.e., by considering only the roots or stems of the words, the comparison between the measurements related data and the reference data can be facilitated. On the one hand, little to no information can be lost by performing word stemming as regards the identification of a measurement. On the other hand, word stemming can increase the homogeneity between the measurement identifiers and the reference data.

For example, the biomarkers "Leucocytes Count" and "Leucocyte Count" are identical apart from the affixes. Thus, by performing word stemming a comparison between the two can yield a full similarity between the two biomarkers.

In some embodiments, the pre-processing step can comprise generating for each of the at least one detected measurement identifier a corresponding measurement identifier data structure. This can facilitate storing and processing the measurement identifier and particularly comparing the measurement identifiers with reference data.

The measurement identifier data structure corresponding to a measurement identifier can be a multiset of the data elements comprised by the measurement identifier. In other words, each data element can be considered as an element of a multiset corresponding to the measurement identifier. The order of the data elements can be disregarded; however, the multiplicity of each data element can be maintained.

A person skilled in the art will appreciate that a multiset (or bag, or mset) is a modification of the concept of a set that, unlike a set, allows for multiple instances for each of its elements.

For example, the measurement identifier data structure corresponding to a measurement identifier can comprise a list data structure, wherein each element in the list data structure comprises a data element of the measurement identifier. In some embodiments, the measurement identifier data structure can comprise for each data element (i.e. for each key) the multiplicity of the data element (i.e. a value). Thus, data elements occurring multiple times in the measurement identifier are represented only once in the list while at the same time their occurrence is maintained.

The list data structure can be unordered. This can be advantageous in instances where two biomarkers are the same, however, the order of the words can be different. For example, "Cholesterol HDL" and "HDL cholesterol" refer to the same measurement, although they comprise a different word order.

Each element in the list data structure comprises a root portion of the data element. A root portion of the data element can be a part or portion of the data element conveying the main information.

For example, the data element can be or can correspond to a word and the root portion of the data element can be or can correspond to the root of the word.

Generating for each of the at least one detected measurement identifier a corresponding measurement identifier data structure can comprise utilizing bag-of-words modeling.

In some embodiments, the pre-processing step can comprise generating for each detected measurement identifier a corresponding data dependent value. The data dependent value corresponding to a measurement identifier can depend on the data comprised by the measurement identifier. Generating for each measurement identifier a data dependent value

can be advantageous as it can facilitate comparing the measurement identifier with the reference data. More particularly, comparing data dependent values (e.g. numbers) can require less computations than directly comparing the measurement identifier with the reference data. Moreover, in some embodiments, the same data dependent value can be generated from two measurement identifiers that can be identical apart from an exchanged order. Thus, by comparing the data dependent values instead of directly comparing the measurement identifiers, identical measurement identifiers apart from an interchange of parts can be detected in a resource-efficient way.

Generating for each of the at least one detected measurement identifier a corresponding data dependent value can comprise executing a data dependent value generating function.

The data dependent value generating function can takes as input a measurement identifier and can output a corresponding data dependent value.

Similarly, the data dependent value generating function can takes as input an identifier references (comprised by the reference data) and can output a corresponding data dependent value.

The data dependent value generating function can be configured to generate similar corresponding data dependent values for similar measurement identifiers. For example, the difference between two measurement identifiers (or between a measurement identifier and an identifier references) can be proportional to a difference between the respective data dependent values.

The data dependent value can be a numerical value. This can be advantageous, as generally comparing numbers can require less computations then comparing strings.

The data dependent value can comprise a hash value corresponding to the measurement identifier.

Executing a data dependent value generating function can comprise executing a hashing function to generate the hash value corresponding to the measurement identifier.

Generating for each of the at least one detected measurement identifier a corresponding data dependent value, can comprise generating a sum of data of the measurement identifier.

For example, generating a sum of data of the measurement identifier can comprise generating a sum of the ASCII values corresponding to the measurement identifier.

The reference data can comprise a plurality of measurement identifier references. For the sake of brevity, the measurement identifier references can be referred to as identifier references. The identifier references can be configured to identify measurements. That is, the plurality of identifier references can comprise a plurality of possible measurement identifiers that can be comprised in the measurement related data. In other words, the reference data can comprise a plurality of measurement identifiers expected to appear in the measurement related data. However, to differentiate the measurement identifiers comprised by the reference data with the measurement identifiers comprised in the measurement related data, the measurement identifiers comprised by the reference data are referred to as measurement identifier reference, or for brevity as identifier reference.

Moreover, each of the sub-steps of the pre-processing step, discussed above mainly in relation with the detected measurement identifiers, can be similarly performed to the identifier references. More particularly, the identifier references can be pre-processed once in an offline process and can be stored in a pre-processed form. Then, the detected measurement identifiers can be pre-processed to bring them in homogeneity with the identifier references. This can facilitate comparing the measurement identifiers with identifier references.

The step of the data processing system comparing the obtained measurement related data with reference data can comprise comparing each detected measurement identifier with the reference data.

Moreover, comparing each detected measurement identifier with the reference data can comprise comparing each detected measurement identifier with at least one identifier reference.

The method can further comprise determining for each detected measurement identifier, respectively, a matching measurement identifier reference based on the respective comparison. That is, the data processing system can compare each detected measurement identifier with the reference data to determine an identifier reference matching to the measurement identifier.

In some embodiments, the method can comprise utilizing at least one heuristic rule to reduce the number of comparisons required to determine for each detected measurement identifier, respectively, a matching measurement identifier reference. As discussed, the reference data can be configured to comprise a plurality of identifier references. Preferably, the plurality of identifier references can include a large number, preferably all, measurement identifiers that can be comprised in the measurement related data. As such, the reference data can be large. Hence, comparing each measurement identifier with the reference data, e.g., with each identifier reference, may require a large number of computations. While this can be inefficient, the present invention can utilize heuristics to reduce the number of computations that may be needed to determine for each detected measurement identifier, respectively, a matching measurement identifier reference. Generally, the heuristic rules aim at comparing the measurement identifier with the most relevant identifier references first. Thus, the likelihood of finding the matching identifier reference without the need of comparing with all the identifier references can be increased. As a result, the number of comparisons, on average, can be reduced as well as the time and computational resources needed to determine for each detected measurement identifier, respectively, a matching measurement identifier reference.

In some embodiments, each comparison between each detected measurement identifier with the reference data can be performed iteratively.

That is, each of the at least one detected measurement identifiers can be compared with the reference data using an iterative process. In other words, each comparison comprises at least one iteration and, in each comparison, a respective detected measurement identifier can be compared with the reference data.

During each iteration the method can comprise determining a set of measurement identifier references. The set of measurement identifier references can be a sub-set of the plurality of measurement identifier references of the reference data.

Moreover, in each iteration the method comprises comparing the respective detected measurement identifier with each of the measurement identifier references comprised by the set of measurement identifier references determined during that iteration. It will be understood that, herein, the respective detected measurement identifier refers to the measurement identifier that is being compared with the reference data during the respective comparison that comprises the iteration.

Heuristic rules can be utilized to determine the set of measurement identifier references in each iteration. The heuristic rules can be configured such that during the first iteration(s) the measurement identifier can be compared with the most relevant identifier references, i.e., with the identifier references comprising the highest likelihood of matching with the measurement identifier.

Determining a set of measurement identifier references during each iteration can comprise selecting the set of measurement identifier references out of the plurality of measurement identifier references comprised by the reference data.

Selecting the set of measurement identifier references out of the plurality of measurement identifier references comprised by the reference data can comprise selecting each measurement identifier reference if the number of data elements of the measurement identifier reference is within a data-element-count range corresponding to that iteration.

The data-element-count range for each iteration can be centered on the number of data elements of the respective detected measurement identifier.

Thus, initially the measurement identifier can be compared with identifier references that comprise a similar number of data elements.

For each comparison between each detected measurement identifier with the reference data the method can comprise extending the data-element-count range during each iteration of the comparison. That is, the upper and/or lower limits of the range are increased/decreased respectively.

In other words, initially, the identifier references with similar number of words to the measurement identifier are considered. Afterwards, if no matching identifier reference is found, in the next iterations identifier references with less similar number of words are considered. During each iteration, the difference between the number of words of the measurement identifier and the number of words of the considered identifier references can be increased.

For each iteration the data-element-count range used during that iteration excludes the data-element-count range used during a previous iteration. That is, the data-element-count range

can be a discontinuous range. This can ensure that each measurement identifier reference is compared only once with the at last one measurement identifier.

The data-element-count range for the first iteration of each comparison between each detected measurement identifier with the reference data can consist of the number of data elements comprised by the detected measurement identifier.

That is, during the first iteration a first set of measurement identifier references can be determined, wherein each measurement identifier reference in the first set of measurement identifier reference comprises the same number of data elements as the respective detected measurement identifier.

Similar to the use of the data element count, the data dependent value can be utilized to select during each iteration the set of identifier references to be compared with the measurement identifier during the respective iteration. The data dependent values can be used alternatively or additionally to the data count values discussed above.

Selecting the set of measurement identifier references out of the plurality of measurement identifier references comprised by the reference data can comprise selecting each measurement identifier reference if the data dependent value of the measurement identifier reference is within a data-dependent-value range corresponding to that iteration.

The data-dependent-value range for each iteration can be centered on the data dependent value of the respective detected measurement identifier.

For each comparison between each detected measurement identifier with the reference data the method can comprise extending the data-dependent-value range during each iteration of the comparison.

That is, the upper and/or lower limits of the range can be increased/decreased respectively.

For each iteration the data-dependent-value range used during that iteration can exclude the data-dependent-value range used during a previous iteration.

That is, the data-element-count range can be a discontinuous range. This can ensure that each measurement identifier reference is compared only once with the at last one measurement identifier.

During each iteration of each comparison between each detected measurement identifier with the reference data the method can comprise calculating a respective similarity metric between the respective detected measurement identifier and each of the measurement identifier references comprised by the set of measurement identifier references determined during that iteration.

In other words, comparing a measurement identifier with an identifier reference can comprise calculating a similarity metric between the measurement identifier and the identifier reference.

The similarity metric calculated between a measurement identifier and a measurement identifier reference can be configured to indicate (e.g. quantify) a similarity between the measurement identifier and the measurement identifier reference.

During each iteration the method can comprise comparing each calculated similarity metric with a matching threshold.

The reference data can comprise the matching threshold.

In some embodiments, the matching threshold can be learned, e.g., during an offline training process.

For each iteration the method can comprise determining whether to execute a next iteration depending on the comparison of each of the calculated similarity metrics with the matching threshold.

For example, if all of the similarity metrics calculated during that iteration are smaller than the matching threshold then the next iteration will be executed.

The method can comprise determining at least one matching measurement identifier reference depending on the comparison of each of the calculated similarity metrics with the matching threshold.

It will be understood, that depending on the respective similarity metrics and the matching threshold, there may be zero, one or a plurality of reference elements.

For each comparison between each detected measurement identifier with the reference data, the method can comprise stopping the comparison when at least one matching measurement identifier reference is determined or when all the measurement identifier references are compared with the respective detected measurement identifier.

In some embodiments, the method can comprise determining for at least one of the detected measurement identifiers a plurality of matching measurement identifier references.

If a plurality of matching measurement identifier references can be determined for a detected measurement identifier, the method can comprise determining only the matching measurement identifier reference that comprises the maximum similarity with the detected measurement identifier as the one corresponding to the detected measurement identifier.

If the measurement identifier can be detected based on the detection of at least one measurement indicator, the method can comprise filtering the plurality of matching measurement identifier references based on the at least one measurement indicator.

Calculating the similarity metric can comprise calculating a Jaccard similarity coefficient between the respective measurement identifier and the respective measurement identifier reference.

In such embodiments, the measurement identifier and the measurement identifier reference can be considered as sets, or multi-sets, of data elements, e.g., as bag of words.

Calculating the similarity metric can comprises calculating a Metaphone distance between a data element of the respective measurement identifier and a data element of the respective measurement identifier reference.

Calculating the similarity metric can comprise calculating a Sørensen–Dice coefficient between the respective measurement identifier and the respective measurement identifier reference.

In such embodiments, the measurement identifier and the measurement identifier reference can be considered as sets, or multi-sets, of data elements, e.g., as bag of words.

The reference data can be configured to identify a plurality of measurements. More particularly, the reference data can comprise a plurality of measurement identifiers (referred to as measurement identifier references).

More particularly, the reference data can comprise for each measurement at least one, preferably a plurality, of measurement identifier reference(s) configured to identify the measurement.

The measurement identifier references configured to identify the same measurement can be associated with each other.

That is the reference data can comprise a plurality of links, each configured to link (i.e. associate) at least two measurement identifier references. Alternatively or additionally, the measurement identifier references can be configured to identify the same measurement can be clustered or grouped together (i.e. can form a cluster).

At least one measurement identifier reference corresponding to a measurement can be configured according to an intermediate code.

The intermediate code can be a standard code. This can allow the present invention to configure the measurement related data according to a standard code.

That is, for each measurement that the reference data is configured to identify, the reference data can comprise at least one measurement identifier reference that can be a standard name or code used to refer to the measurement.

The intermediate code can be configured according to the Logical Observation Identifiers Names and Codes (LOINC) database. In other words, the intermediate code can utilize the nomenclature of the LOINC database.

The reference data can comprise for each measurement, in addition to the at least one measurement identifier reference configured according to an intermediate code, at least one further measurement identifier reference, wherein the at least one further measurement identifier reference can be configured according to a code different from the intermediate code.

For each measurement, each further measurement identifier reference in addition to the at least one measurement identifier reference configured according to an intermediate code, can be associated with the at least one measurement identifier reference configured according to an intermediate code.

Thus, two or more measurement identifier references can be linked or associated with each other if they are linked or associated to the same measurement identifier reference configured according to an intermediate code.

For example, for each measurement the reference data can comprise a corresponding cluster of identifier references.

The reference data can comprise, associated to each of the at least one measurement identifier references, a corresponding code specifier. Each corresponding code specifier can be configured to specify the code that the respective measurement identifier reference corresponds to.

Moreover, the reference data can be configured to characterize a plurality of measurements.

The reference data can comprise for each measurement at least one, preferably a plurality, of reference characteristic(s).

The at least one reference characteristic can comprise at least one of an object, component, substance, sample or specimen to be measured, a unit of measurement, an interval of time over which a measurement is to be made, a scale type (i.e. a measured value type, e.g., a numerical value or text value), and a classification of how the measurement is to be made.

Configuring the measurement related data according to the second code can comprise determining for each measurement identifier a replacing measurement identifier reference and replacing each measurement identifier with the respective replacing measurement identifier reference.

The replacing measurement identifier reference can be a measurement identifier reference configured according to the second code.

Determining for each measurement identifier a replacing measurement identifier reference can depend on the respective matching measurement identifier reference determined for the respective measurement identifier. For example, the matching identifier reference and the replacing measurement identifier reference can be linked, i.e., can correspond to the same measurement.

Determining for each measurement identifier a replacing measurement identifier reference can comprise determining whether the respective matching measurement identifier reference determined for the respective measurement identifier can be configured according to the second code.

That is, the present method can first determine for each measurement identifier in the measurement related data a matching identifier reference. As discussed, the matching identifier reference can be an identifier reference that comprises a high similarity with the measurement identifier. In other words, the present method can determine a respective identifier reference that matches with each measurement identifier.

In some embodiments, the method can comprise outputting the matching identifier reference.

In some embodiments, the method can comprise outputting the matching identifier reference with corresponding reference characteristics that can be comprised by the reference data.

In some instances, the matching identifier reference can correspond to the second code. In such cases, the matching identifier reference can be at the same time the replacing identifier reference. However, in some instances the matching identifier reference may not correspond to the second code. For example, the matching identifier reference can correspond to the first code. In such cases, the replacing identifier reference can be determined to be an identifier reference that is linked with the matching identifier reference and that corresponds to the second code.

That is, in some embodiments, configuring the measurement related data according to the second code can be a two-tier process. First a matching identifier reference can be determined for each measurement identifier. This can facilitate determining for each measurement identifier, which measurement it identifies. As discussed, the reference data can comprise for each measurement a plurality of identifier references. Each identifier reference can correspond to a respective code. Thus, determining a matching identifier reference for each measurement

identifier, allows for configuring the measurement related data according to any code for which the reference data comprise identifier references. Thus, in the second tier the replacing identifier reference can be determined based on the matching identifier reference and the second code.

The method can comprise outputting the replacing identifier reference. In some embodiments, the method can comprise outputting the matching identifier reference with corresponding reference characteristics that can be comprised by the reference data.

The method can comprise utilizing the code specifier corresponding to the matching measurement identifier reference to determine whether the respective matching measurement identifier reference determined for the respective measurement identifier is configured according to the second code.

Determining for each measurement identifier a replacing measurement identifier reference can comprise determining the respective replacing measurement identifier reference for each measurement identifier, to be the respective matching measurement identifier reference if the matching measurement identifier reference is determined to be configured according to the second code.

The method can comprise for each detected measurement identifier finding a measurement identifier reference that is configured according to the second code and that is associated with the respective matching measurement identifier reference, if the respective matching measurement identifier reference is not configured according to the second code.

Determining for each measurement identifier a replacing measurement identifier reference can comprise determining the respective replacing measurement identifier reference, for each measurement identifier, to be the measurement identifier reference that is configured according to the second code and that is associated with the respective matching measurement identifier reference.

In some instances, the first code can be different from the intermediate code and the second code can be the intermediate code. That is, the method can comprise configuring the measurement related data from an arbitrary code to the intermediate code, e.g., to a standard code.

In some instances, the first code can be the intermediate code and the second code can be different from the intermediate code. That is, the method can comprise configuring the measurement related data from an intermediate code, e.g., a standard code, to an arbitrary code.

In some instances, the first code can be different from the intermediate code and the second code can be different from the intermediate code. That is, the method can comprise configuring the measurement related data from a first arbitrary code to a second arbitrary code without utilizing any intermediate code.

Again, in some embodiments, the reference data can comprise for each measurement a plurality of identifier references. Each reference identifier can be a measurement identifier corresponding to a respective code. Thus, for each measurement the reference data can comprise identifiers used by different codes. The identifier references corresponding to the same measurement can be clustered together. The present method can match the detected measurement identifiers with a respective cluster by determining the matching identifier reference and then determining to which cluster the matching identifier reference corresponds to. It can be advantageous for each cluster to comprise a plurality of identifier references, each corresponding to a respective code. Thus, the measurement related data can be configured according to an arbitrary code and the respective cluster (i.e. a matching identifier reference) can be determined with a high likelihood. Typically, the richer the reference data (i.e. the more identifier references they comprise) the higher the likelihood of matching the measurement identifier with identifier references.

Once a measurement identifier can be matched with a cluster, it can be replaced by any identifier reference in that cluster. In other words, the measurement identifier and any identifier reference corresponding to the matched cluster, can refer to the same measurement. The matched cluster refers to the cluster that the matching identifier reference corresponds to. Thus, the measurement related data can be configured according to different codes by simply replacing each measurement identifier with the appropriate identifier reference from the matched cluster.

The method can comprise the data processing system outputting the measurement related data configured according to the second code.

The method can comprise a sending node generating the measurement related data.

For example, the sending node can be a device/system which can be programmed to generate data configured according to the first code.

The method can comprise communicating the measurement related data from the sending node to the data processing system.

The method can comprises communicating the measurement related data from the sending node to the data processing system through an electronic data communication network.

That is, the sending node can be interconnected with the data processing system through an electronic data communication network.

The method can comprise a receiving node receiving the measurement related data configured according to the second code.

For example, the receiving node can be a device/system which can be programmed to read and/or "understand", i.e., decode, data configured according to the second code.

The method can comprise communicating the measurement related data from the data processing system to the receiving node.

The method can comprise communicating the measurement related data from the data processing system to the receiving node through an electronic data communication network.

That is, the receiving node can be interconnected with the data processing system through an electronic data communication network.

The measurement related data can comprise measurement instruction data. In such embodiments, the sending node can be a measurement requesting node. Moreover, the receiving node can be a measurement performing node.

In some embodiments, the measurement related data can comprise measurement result data. In such embodiments, the sending node can be a measurement performing node. Moreover, the receiving node can be a measurement requesting node.

The method can comprise a measurement requesting node generating measurement instruction data configured according to the first code and sending the measurement instruction data to the data processing system. The data processing system can send the measurement instruction data configured according to the second code to a measurement performing node. The measurement performing node can perform the requested measurement(s) and can generate measurement result data configured according to the second code. Moreover, the measurement performing node can send the measurement result data to the data processing system. The data processing system can configure the measurement result data according to the first code and can send the measurement result data to the measurement requesting node.

Thus, the communication between the measurement requesting node and the measurement performing node can be facilitated by the data processing system.

It will be understood that configuring the measurement related data from the first code to the second code can be equivalent (in that it comprises similar steps) to configuring the measurement related data from the second code to the first code.

The method can be a computer-implemented method.

In some embodiments, the method can be carried out by the data processing system.

In a further aspect, the present invention relates to a data processing system comprising an input unit configured to obtain measurement related data configured according to a first code and a matching unit configured to compare the obtained measurement related data with reference data and to configure the measurement related data according to a second code based on the comparison.

The input unit comprises a network interface.

The data processing system can further comprise an entities recognition unit configured to detect in the obtained measurement related data at least one of a data element, a data portion, a measurement identifier and a measurement indicator.

The entities recognition unit can be configured to carry out any of the detection steps of the method discussed above.

The entities recognition unit can be configured to obtain the measurement related data from the input unit.

The data processing system can comprise an online pre-processing unit configured to pre-process the at least one detected measurement identifier.

The entities recognition unit can be configured to detect at least one measurement identifier and to provide the measurement identifier to the online pre-processing unit.

The online pre-processing unit can be configured to carry out any of the pre-processing steps of the method discussed above to pre-process the measurement identifiers.

The matching unit can be configured to carry out any of the comparison steps of the method discussed above.

The data processing system can comprise a similarity metric calculating unit which can be configured to receive two inputs and to calculate a similarity metric between the first input and the second input.

The matching unit can be configured to utilize the similarity metric calculating unit to calculate the similarity metric between a measurement identifier and an identifier reference.

The data processing system can comprise an offline pre-processing unit which can be configured to pre-process the reference data.

The reference data can comprise a plurality of identifier references and the offline pre-processing unit can be configured to pre-process the identifier references.

The offline pre-processing unit can be configured to carry out the same pre-processing steps as the online pre-processing unit to pre-process the identifier references.

The data processing system can comprise an output unit which can be configured to output the measurement related data configured according to the second code.

The output unit can comprise at least one of a display, a printer, a fax, and a network card.

The data processing system can comprise a learning unit configured to extend the reference data.

The data processing system can be configured to carry out the method according to any of the method embodiments.

It will be understood that the data processing system and the method discussed above can comprise corresponding features. In other words, features and advantages discussed above with respect to the method can also be valid for the data processing system. For the sake of brevity, they may not be repeated in the following.

In a further aspect, the present invention relates to a communication system comprising a data processing system wherein the data processing system is configured to obtain measurement related data configured according to a first code, compare the obtained measurement related data with reference data and configure the measurement related data (10, 20) according to a second code based on the comparison.

It will be understood that the communication system and the method discussed above can comprise corresponding features. In other words, features and advantages discussed above with respect to the method can also be valid for the communication system. For the sake of brevity, they may not be repeated in the following.

The communication system can comprise a memory component configured to store the reference data.

The data processing system can be configured to access the memory component.

The memory component can be integrated into the data processing system.

The communication system can further comprise a sending node configured to generate the measurement related data according to the first code. For example, the sending node can be configured to store, process and/or communicate data configured according to the first code. For example, the sending node can encode and decode data based on the first code.

The sending node can be configured to electronically communicate with the data processing system.

The communication system can further comprise a receiving node configured to receive the measurement related data configured according to the second code. For example, the receiving node can be configured to store, process and/or communicate data configured according to the second code. For example, the receiving node can encode and decode data based on the second code.

The receiving node can be configured to electronically communicate with the data processing system.

The communication system can be configured to carry out the method according to any of the preceding method embodiments.

The data processing system can be configured to carry out the method according to any of the preceding method embodiments to configure the measurement related data (generated by the sending node) according to the second code.

The second code can be an intermediate code. For example, the data processing system can be utilized to configure the measurement related data (generated by the first node) according to a standard code.

In some embodiments, the data processing system can be provided as an extension to the first node. For example, the data processing system can be integrated into the first node.

In some embodiments, the data processing system can be provided as an extension to the second node. For example, the data processing system can be integrated into the second node.

In general, the data processing system can be provided as an extension and/or can be integrated into any node in a communication network. This can facilitate interfacing the node with other nodes in the communication network.

In some embodiments, the data processing system can be configured to carry out the method according to any of the preceding method embodiments to facilitate the communication between the sending node and the receiving node.

In a further aspect, the present invention relates to a computer program product comprising instructions, which when the program is executed by a computer can cause the computer to carry out the method according to any of the preceding method embodiments. The computer can comprise the data processing system.

In a further aspect, the present invention relates to a computer-readable storage medium comprising instructions, which when the instructions are executed by a computer can cause the computer to carry out the method according to any of the preceding method embodiments. The computer can comprise the data processing system.

In a further aspect, the present invention relates to a use of the method and/or system according to any of the preceding method and/or system embodiments for configuring measurement related data from a first code to a second code.

In a further aspect, the present invention relates to a use of the method and/or system according to any of the preceding method and/or system embodiments for configuring measurement related data from a first code to a second code, wherein the second code is an intermediate code, such as, a standard code.

In a further aspect, the present invention relates to a use of the method and/or system according to any of the preceding method and/or system embodiments for facilitating a communication between a sending node configured to generate (e.g. encode) measurement related data according to a first code and a receiving node configured to receive (e.g., decode) measurement related data according to a second code.

In a further aspect, the present invention relates to a use of the method and/or system according to any of the preceding method and/or system embodiments for facilitating a transmission of measurement related data from a sending node configured to generate (e.g. encode) measurement related data according to a first code to a receiving node configured to receive (e.g. decode) measurement related data according to a second code.

The present technology is also defined by the following numbered embodiments.

Below, numbered method embodiments will be discussed. When reference is herein made to method embodiments, these embodiments are meant.

1. A method comprising:
 - a data processing system (40) obtaining measurement related data (10, 20) configured according to a first code (70A);
 - the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50);
 - based on the comparison, the data processing system (40) configuring the measurement related data (10, 20) according to a second code (70B).
2. The method according to any of the preceding embodiments, wherein the obtained measurement related data (10, 20) comprises at least one measurement identifier (3) configured to identify a measurement.

For example, the measurement identifier (3) can be a name of a measurement (e.g. a biomarker name).

3. The method according to the preceding embodiment, wherein the method comprises detecting at least one measurement identifier (3) in the obtained measurement related data (10, 20).

Preferably, the method can comprise detecting each of the at least one measurement identifiers (3) comprised in the obtained measurement related data (10, 20).

4. The method according to any of the preceding embodiments, wherein the obtained measurement related data (10, 20) comprises a plurality of data portions (30).
5. The method according to the preceding embodiment, wherein the obtained measurement related data (10, 20) consists of a plurality of data portions (30).
6. The method according to any of the 2 preceding embodiments wherein the data portions (30) are non-intersecting portions of the obtained measurement related data (10, 20).

For example, the data portions (30) can be blocks of data within the measurement related data (10, 20), i.e., the measurement related data (10, 20) can be arranged in blocks of data (or data blocks).

7. The method according to any of the 3 preceding embodiments, wherein the method comprises detecting at least one data portion (30) in the obtained measurement related data (10, 20).

Preferably, the method can comprise detecting each of the at least one data portion (30) comprised in the obtained measurement related data (10, 20).

8. The method according to any of the 4 preceding embodiments, wherein the data portions (30) correspond to respective lines of the obtained measurement related data (10, 20).

9. The method according to any the preceding embodiments, wherein the obtained measurement related data (10, 20) comprises portion delimiters (33) configured to specify the boundaries of the data portions (30).

That is, the data portions (30) can be separated from each other using portion delimiters (33).

10. The method according to the preceding embodiment, wherein the portion delimiters (33) are new line characters (33).

11. The method according to any of the preceding embodiments and with the features of embodiment 3 and 4, wherein detecting at least one measurement identifier (3) comprises detecting at least one data portion (30) of the obtained measurement related data (10, 20) that comprises a measurement identifier (3).

12. The method according to the preceding embodiment, wherein detecting at least one data portion (30) of the obtained measurement related data (10, 20) that comprises a measurement identifier (3) comprises detecting each data portion (30) of the obtained measurement related data (10, 20).

13. The method according to the preceding embodiment, wherein detecting at least one data portion (30) of the obtained measurement related data (10, 20) that comprises a measurement identifier (3) comprises

determining for each of the detected data portions (30) whether it comprises a measurement identifier (3).

14. The method according to any of the preceding embodiments, wherein the measurement related data (10, 20) comprises a plurality of data elements (35).

15. The method according to the preceding embodiment and with the features of embodiment 4, wherein each data portion (30) comprises at least one data element (35).

16. The method according to any of the 2 preceding embodiments, wherein each data element (35) comprises a plurality of data bits.

17. The method according to any of the 3 preceding embodiments, wherein each data element (35) comprises at least one byte of data.

18. The method according to any of the 4 preceding embodiments, wherein each data element (35) comprises at least one character.

For example, each data element can comprise at least one ASCII character.

19. The method according to any of the 5 preceding embodiments and with the features of embodiment 2, wherein each of the at least one measurement identifiers (3) comprises at least one data element (35).

20. The method according to the preceding embodiment and with the features of embodiment 3, wherein detecting at least one measurement identifier (3) comprises

determining for each data element (35) in the measurement related data (10, 20) whether it corresponds to a measurement identifier (3).

21. The method according to the preceding embodiment, wherein determining for each data element (35) in the measurement related data (10, 20) whether it corresponds to a measurement identifier (3)

comprises comparing each data element (35) and/or a sequence of data elements (35) with a plurality of measurement identifier references (53).

22. The method according to the preceding embodiment, wherein the reference data (50) comprise the plurality of measurement identifier references (53).

23. The method according to any of the preceding embodiments and with the features of embodiment 14, wherein the data elements (35) are identifiable.

24. The method according to any the preceding embodiments, wherein the measurement related data (10, 20) comprises element delimiters (37) configured to specify boundaries of the data elements (35).

25. The method according to any of the preceding embodiments, wherein the method comprises detecting at least one measurement indicator (2) in the obtained measurement related data (10, 20).

26. The method according to the preceding embodiment, wherein the at least one measurement indicator (2) comprises at least one of
numerical values, ordinal values, nominal values, qualitative data, quantitative data, unit of measurement, range and range specifier.

27. The method according to any of the 2 preceding embodiments and with the features of embodiment 3, wherein detecting at least one measurement identifier (3) comprises
detecting at least one measurement identifier (3) based on the detection of at least one measurement indicator (2).

The above can be based on the rationale that measurement identifiers can typically be associated with numbers, units and ranges for specifying the measurement in more detail.

28. The method according to any of the 3 preceding embodiments, wherein the method comprises determining the location of the at least one measurement indicator (2) in the obtained measurement related data (10, 20).

29. The method according to the preceding embodiment and with the features of embodiment 3, wherein detecting at least one measurement identifier (3) comprises detecting at least one measurement identifier (3) based on the location of the at least one measurement indicator (2) in the obtained measurement related data (10, 20).

For example, data that precedes or follows the at least one measurement indicator (2) can be determined (or hypothesized with a respective certainty measure, e.g., likelihood) to be a measurement identifier (3).

30. The method according to any of the 5 preceding embodiments and with the features of embodiment 4, wherein the method comprises determining whether the at least one data portion (30) comprises at least one measurement indicator (2).

31. The method according to the preceding embodiment and with the features of embodiment 11, wherein detecting at least one data portion (30) of the obtained measurement related data (10, 20) that comprises a measurement identifier (3) comprises determining that at least one data portion (30) comprises a measurement identifier (3) if the at least one data portion (30) comprises at least one measurement indicator (2).

32. The method according to the preceding embodiment, wherein the method comprises upon detecting at least one measurement indicator (2) in a data portion (30), determining that the remaining data in the data portion (30) comprise the measurement identifier (3).

33. The method according to any of the 2 preceding embodiments, wherein the method comprises upon detecting at least one measurement indicator (2) in a data portion (30), determining that the remaining data in the data portion (30) is the measurement identifier (3).

34. The method according to any of the preceding embodiments and with the features of embodiment 25, wherein the at least one measurement indicator (2) comprises a characteristic type.

35. The method according to the preceding embodiment, wherein the characteristic type of the at least one measurement indicator (2) comprises at least one of
numerical value, ordinal value, nominal value, qualitative data, quantitative data, unit of measurement, range and range specifier.

36. The method according to any of the 2 preceding embodiments, wherein detecting at least one measurement indicator (2) in the obtained measurement related data (10, 20) comprises
determining the characteristic type of the at least one measurement indicator (2).

37. The method according to any of the preceding embodiments and with the features of embodiment 25, wherein detecting at least one measurement indicator (2) in the obtained measurement related data (10, 20) comprises using regular expressions.

The regular expressions can be used to detect (e.g. search for) the at least one measurement indicator (2). Alternatively or additionally, the regular expressions can be used to detect data portions that can comprise at least one measurement identifier

38. The method according to any of the preceding embodiments and with the features of embodiment 25, wherein detecting at least one measurement indicator (2) in the obtained measurement related data (10, 20) comprises using heuristic rules.

39. The method according to any of the preceding embodiments and with the features of embodiment 25, wherein detecting at least one measurement indicator (2) comprises using a string-searching algorithm.

40. The method according to any of the preceding embodiments and with the features of embodiment 25,

wherein each of the at least one measurement indicators (2) comprises at least one data element (35) and

wherein detecting at least one measurement indicator (2) comprises determining for each data element (35) in the measurement related data (10, 20) whether it corresponds to a measurement indicator (2).

41. The method according to the preceding embodiment, wherein determining for each data element (35) in the measurement related data (10, 20) whether it corresponds to a measurement indicator (2) comprises

comparing each data element (35) and/or a sequence of data elements (35) with a plurality of reference characteristics (55).

42. The method according to the preceding embodiment, wherein the reference data (50) comprises the plurality of reference characteristics (55).

43. The method according to any of the preceding embodiments and with the features of embodiment 3, wherein the method comprises pre-processing the at least one detected measurement identifier (3)

to facilitate the step of the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50).

44. The method according to the preceding embodiment wherein the pre-processing step is performed before the step of the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50).

45. The method according to any of the 2 preceding embodiments, wherein the data processing system (40) performs the pre-processing step.

46. The method according to any of the 2 preceding embodiments, wherein the pre-processing step comprises performing a data cleaning of the at least one detected measurement identifier (3).

47. The method according to the preceding embodiment and with features of embodiment 19, wherein the data cleaning step comprises detecting at least one data element (35'), comprised by the at least one detected measurement identifier (3), which does not facilitate the identification of a measurement.

The referral number (35') is used to refer to data elements (35) comprised by the measurement identifiers (3) which do not facilitate the identification of a measurement. The data elements (35') which do not facilitate the identification of a measurement may also be referred to as irrelevant data elements (35').

48. The method according to the preceding embodiment, wherein detecting at least one data element (35') which does not facilitate the identification of a measurement comprises providing a data cleaning database (510) and determining for each data element (35, 35') comprised by the at least one detected measurement identifier (3) whether it is part of the data cleaning database (510).

49. The method according to the preceding embodiment, wherein the reference data (50) comprise the data cleaning database (510).

50. The method according to the preceding embodiments, wherein the data cleaning database (510) comprises a plurality of stop-words, symbols and punctuation marks.

51. The method according to any of the 4 preceding embodiments, wherein the method comprises the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50) without utilizing the at least one data element (35') which does not facilitate the identification of a measurement.

52. The method according to any of the 5 preceding embodiments, wherein the pre-processing step comprises removing from the measurement identifiers (3) the at least one data element (35') which does not facilitate the identification of a measurement.

53. The method according to any of the preceding embodiments and with features of embodiments 19 and 43, wherein pre-processing the at least one detected measurement identifier (3) comprises counting the number of data elements (35) comprised by the at least one detected measurement identifier (3).

Thus, a data element count can be determined for some or each of the at least one detected measurement identifier (3).

54. The method according to the preceding embodiment and with features of embodiment 47, wherein counting the number of data elements (35) comprised by the at least one detected measurement identifier (3) comprises skipping the detected data elements (35') which do not facilitate the identification of a measurement.

55. The method according to any of the 2 preceding embodiments and with the features of embodiment 52, wherein counting the number of data elements (35) comprised by the at least one detected measurement identifier (3) is performed after

removing the data elements (35') which do not facilitate the identification of a measurement from the measurement identifiers (3).

56. The method according to any of the preceding embodiments and with features of embodiments 19 and 43, wherein the pre-processing step comprises

replacing at least one data element (35) comprised by the at least one detected measurement identifier (3) with an equivalent data element (35E).

57. The method according to the preceding embodiment, wherein replacing at least one data element (35) comprised by the at least one detected measurement identifier (3) with an equivalent data element (35E) comprises

utilizing an equivalent data elements database (540),

wherein the equivalent data elements database (540) comprises a plurality of data elements (35) each associated with at least one equivalent data element (35E).

58. The method according to the preceding embodiment, wherein the equivalent data elements database (540) comprises a synonyms dictionary.

59. The method according to any of the 2 preceding embodiments, wherein utilizing an equivalent data elements database (540) comprises

searching the equivalent data elements database (540) for at least one data element (35) comprised by the at least one detected measurement identifier (3).

60. The method according to any of the 4 preceding embodiments, wherein the data element (35) and the equivalent data element (35E) convey same or similar information.

It will be understood that each data element (35) of the measurement related data (10, 20) can be configured to convey information.

61. The method according to any of the 5 preceding embodiments, wherein the data element (35) and the equivalent data element (35E) are synonyms.

62. The method according to any of the 6 preceding embodiments, wherein the data element (35) and the equivalent data element (35E) comprise different grammatical forms of a same word.
63. The method according to any of the 7 preceding embodiments, wherein the equivalent data element (35E) comprises a word stem of the data element (35).
64. The method according to any of the 8 preceding embodiments, wherein the equivalent data element (35E) comprises a word root of the data element (35).
65. The method according to any of the 9 preceding embodiments, wherein the pre-processing step comprises performing word stemming on the data elements (35).
66. The method according to any of the preceding embodiments and with features of embodiment 43, wherein the pre-processing step comprises
generating for each of the at least one detected measurement identifier (3) a corresponding measurement identifier data structure (5).
67. The method according to the preceding embodiment and with features of embodiment 19, wherein the measurement identifier data structure (5) corresponding to a measurement identifier (3) is a multiset of the data elements (35) comprised by the measurement identifier (5).
68. The method according to any of the 2 preceding embodiments and with features of embodiment 19, wherein the measurement identifier data structure (5) corresponding to a measurement identifier (3) comprises
a list data structure, wherein each element in the list data structure comprises a data element (35) of the measurement identifier (3).
69. The method according to the preceding embodiment, wherein the list data structure is unordered.
70. The method according to any of the 2 preceding embodiments, wherein each element in the list data structure comprises a root portion of the data element (35).

For example, the data element (35) can be or can correspond to a word and the root portion of the data element (35) can be or can correspond to the root of the word.

71. The method according to any of the 5 preceding embodiments, wherein generating for each of the at least one detected measurement identifier (3) a corresponding measurement identifier data structure (5) comprises
utilizing bag-of-words modeling.

72. The method according to any of the preceding embodiments and with features of embodiment 43, wherein the pre-processing step comprises
generating for each detected measurement identifier (3) a corresponding data dependent value (7),
wherein the data dependent value (7) corresponding to a measurement identifier (3) depends on the data comprised by the measurement identifier (3).

73. The method according to the preceding embodiment, wherein generating for each of the at least one detected measurement identifier (3) a corresponding data dependent value (7) comprises
executing a data dependent value generating function,
wherein the data dependent value generating function takes as input a measurement identifier (3) and outputs a corresponding data dependent value (7).

74. The method according to the preceding embodiment, wherein the data dependent value generating function is configured to generate similar corresponding data dependent values (7) for similar measurement identifiers (3).

75. The method according to any of the 3 preceding embodiments, wherein the data dependent value (7) is a numerical value.

76. The method according to any of the 4 preceding embodiments, wherein the data dependent value (7) comprises a hash value (7) corresponding to the measurement identifier (3).

77. The method according to the preceding embodiment and with the features of embodiment 73, wherein executing a data dependent value generating function comprises

executing a hashing function to generate the hash value (7) corresponding to the measurement identifier (3).

78. The method according to any of the 6 preceding embodiments, wherein generating for each of the at least one detected measurement identifier (3) a corresponding data dependent value (7), comprises

generating a sum of data of the measurement identifier (3).

79. The method according to the preceding embodiment, wherein generating a sum of data of the measurement identifier (3) comprises

generating a sum of the ASCII values corresponding to the measurement identifier (3).

80. The method according to any of the preceding embodiments, wherein the reference data (50) comprises a plurality of measurement identifier references (53).

81. The method according to the preceding embodiment and with the features of embodiment 3, wherein the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50) comprises

comparing each detected measurement identifier (3) with the reference data (50).

82. The method according to the preceding embodiment, wherein the method comprises determining for each detected measurement identifier (3), respectively, a matching measurement identifier reference (53M) based on the comparison between the measurement identifier (3) and the reference data (50).

83. The method according to the preceding embodiment, wherein the method comprises utilizing at least one heuristic rule to reduce the number of comparisons required to determine for each detected measurement identifier (3), respectively, a matching measurement identifier reference (53M).

84. The method according to any of the 3 preceding embodiments, wherein each comparison between each detected measurement identifier (3) with the reference data (50) is performed iteratively.

That is, each of the at least one detected measurement identifiers (3) is compared with the reference data (50) using an iterative process. In other words, each comparison

comprises at least one iteration and, in each comparison, a respective detected measurement identifier (3) is compared with the reference data (50).

85. The method according to the preceding embodiment, wherein during each iteration of each comparison between each detected measurement identifier (3) with the reference data (50) the method comprises

determining a set of measurement identifier references (53), wherein the set of measurement identifier references (53) is a sub-set of the plurality of measurement identifier references (53) of the reference data (50) and

comparing the respective detected measurement identifier (3) with each of the measurement identifier references (53) comprised by the set of measurement identifier references (53) determined during that iteration.

It will be understood that the respective detected measurement identifier (3) refers to the measurement identifier (3) that is being compared with the reference data (50) during the respective comparison that comprises the iteration.

86. The method according to the preceding embodiment, wherein determining a set of measurement identifier references (53) during each iteration comprises selecting the set of measurement identifier references (53) out of the plurality of measurement identifier references (53) comprised by the reference data (50).

87. The method according to the preceding embodiment and with the features of embodiment 53, wherein selecting the set of measurement identifier references (53) out of the plurality of measurement identifier references (53) comprised by the reference data (50) comprises

selecting each measurement identifier reference (53) if the number of data elements of the measurement identifier reference (53) is within a data-element-count range corresponding to that iteration.

88. The method according to the preceding embodiment, wherein the data-element-count range for each iteration is centered on the number of data elements (35) of the respective detected measurement identifier (3).

89. The method according to any of the 2 preceding embodiments, wherein for each comparison between each detected measurement identifier (3) with the reference data (50) the method comprises

extending the data-element-count range during each iteration of the comparison.

That is, the upper and/or lower limits of the range are increased/decreased respectively.

90. The method according to the preceding embodiment, wherein for each iteration the data-element-count range used during that iteration excludes the data-element-count range used during a previous iteration.

That is, the data-element-count range can be a discontinuous range. This can ensure that each measurement identifier reference is compared only once with the at least one measurement identifier.

91. The method according to any of the 4 preceding embodiments, wherein the data-element-count range for the first iteration of each comparison between each detected measurement identifier (3) with the reference data (50) consists of the number of data elements (35) comprised by the detected measurement identifier (3).

That is, during the first iteration a first set of measurement identifier references can be determined, wherein each measurement identifier reference in the first set of measurement identifier reference comprises the same number of data elements as the respective detected measurement identifier (3).

92. The method according to any of the 6 preceding embodiments and with the features of embodiment 72, wherein selecting the set of measurement identifier references (53) out of the plurality of measurement identifier references (53) comprised by the reference data (50) comprises

selecting each measurement identifier reference (53) if the data dependent value (7) of the measurement identifier reference (53) is within a data-dependent-value range corresponding to that iteration.

93. The method according to the preceding embodiment, wherein data-dependent-value range for each iteration is centered on the data dependent value (7) of the respective detected measurement identifier (3).

94. The method according to any of the 2 preceding embodiments, wherein for each comparison between each detected measurement identifier (3) with the reference data (50) the method comprises

extending the data-dependent-value range during each iteration of the comparison.

That is, the upper and/or lower limits of the range are increased/decreased respectively.

95. The method according to the preceding embodiment, wherein for each iteration the data-dependent-value range used during that iteration excludes the data-dependent-value range used during a previous iteration.

That is, the data-element-count range can be a discontinuous range. This can ensure that each measurement identifier reference is compared only once with the at last one measurement identifier.

96. The method according to any of the preceding embodiments and with the features of embodiment 84, wherein during each iteration of each comparison between each detected measurement identifier (3) with the reference data (50) the method comprises

calculating a respective similarity metric between the respective detected measurement identifier (3) and each of the measurement identifier references (53) comprised by the set of measurement identifier references (53) determined during that iteration.

The similarity metric calculated between a measurement identifier (3) and a measurement identifier reference (53) can be configured to indicate (e.g. quantify) a similarity between the measurement identifier (3) and the measurement identifier reference (53).

97. The method according to the preceding embodiment, wherein during each iteration the method comprises

comparing each calculated similarity metric with a matching threshold.

98. The method according to the preceding embodiment, wherein the reference data (50) comprises the matching threshold.

99. The method according to any of the 2 preceding embodiments, wherein for each iteration the method comprises determining whether to execute a next iteration depending on the comparison of each of the calculated similarity metrics with the matching threshold.

For example, if all of the similarity metrics calculated during that iteration are smaller than the matching threshold then the next iteration will be executed.

100. The method according to any of the 3 preceding embodiments and with the features of embodiment 82, wherein the method comprises determining at least one matching measurement identifier reference (53M) depending on the comparison of each of the calculated similarity metrics with the matching threshold.

It will be understood, that depending on the respective similarity metrics and the matching threshold, there may be zero, one or a plurality of reference elements.

101. The method according to any of the 4 preceding embodiments, for each comparison between each detected measurement identifier (3) with the reference data (50) the method comprises stopping the comparison

when at least one matching measurement identifier reference (53M) is determined or when all the measurement identifier references (53) are compared with the respective detected measurement identifier (3).

102. The method according to any of the preceding embodiments and with the features of embodiment 82, wherein the method comprises determining for at least one of the detected measurement identifiers (3) a plurality of matching measurement identifier references (53M).

103. The method according to the preceding embodiment and with the features of embodiment 100, wherein if a plurality of matching measurement identifier references (53M) are determined for a detected measurement identifier (3) the method comprises

determining only the matching measurement identifier reference (53M) that comprises the maximum similarity with the detected measurement identifier (3) as the one corresponding to the detected measurement identifier (3).

104. The method according to any of the 2 preceding embodiments and with the features of embodiment 27, wherein if the measurement identifier (3) is detected based on the detection of at least one measurement indicator (2) the method comprises

filtering the plurality of matching measurement identifier references (53M) based on the at least one measurement indicator (2).

105. The method according to any of the preceding embodiments and with the features of embodiment 96, wherein calculating the similarity metric comprises calculating a Jaccard similarity coefficient between the respective measurement identifier (3) and the respective measurement identifier reference (53).

In such embodiments, the measurement identifier (3) and the measurement identifier reference (53) can be considered as sets or as multi-sets of data elements (35), e.g., as bag of words.

106. The method according to any of the preceding embodiments and with the features of embodiments 14 and 96, wherein calculating the similarity metric comprises calculating a Metaphone distance between a data element (35) of the respective measurement identifier (3) and a data element (35) of the respective measurement identifier reference (53).

107. The method according to any of the preceding embodiments and with the features of embodiment 96, wherein calculating the similarity metric comprises calculating a Sørensen–Dice coefficient between the respective measurement identifier (3) and the respective measurement identifier reference (53).

In such embodiments, the measurement identifier (3) and the measurement identifier reference (53) can be considered as sets or as multi-sets of data elements (35), e.g., as bag of words.

108. The method according to any of the preceding embodiments, wherein the reference data (50) are configured to identify a plurality of measurements.

109. The method according to the preceding embodiment and with the features of embodiment 80, wherein the reference data (50) comprise for each measurement at least one, preferably a plurality, of measurement identifier reference(s) (53) configured to identify the measurement.

110. The method according to the preceding embodiment, wherein the measurement identifier references (53) configured to identify the same measurement are associated with each other.

That is the reference data (50) can comprise a plurality of links, each configured to link (i.e. associate) at least two measurement identifier references (53). Alternatively or additionally, the measurement identifier references (53) configured to identify the same measurement can be clustered or grouped together (i.e. can form a cluster).

111. The method according to the preceding embodiment, wherein at least one measurement identifier reference (53) corresponding to a measurement is configured according to an intermediate code (70I).

112. The method according to the preceding embodiment, wherein the intermediate code (70I) is a standard code.

That is, for each measurement that the reference data (50) is configured to identify, the reference data (50) can comprise at least one measurement identifier reference (53) that can be a standard name or code used to refer to the measurement.

113. The method according to any of the 2 preceding embodiments, wherein the intermediate code (70I) is configured according to the Logical Observation Identifiers Names and Codes (LOINC) database.

114. The method according to any of the 3 preceding embodiments, wherein the reference data (50) comprises for each measurement, in addition to the at least one measurement identifier reference (53) configured according to an intermediate code (70I), at least one further measurement identifier reference (53),

wherein the at least one further measurement identifier reference (53) is configured according to a code (70) different from the intermediate code (70I).

115. The method according to the preceding embodiment, wherein for each measurement, each further measurement identifier reference (53) in addition to the at least one measurement identifier reference (53) configured according to an intermediate code (70I), is

associated with the at least one measurement identifier reference (53) configured according to an intermediate code (70I).

Thus, two or more measurement identifier references (53) can be linked or associated with each other if they are linked or associated to the same measurement identifier reference (53) configured according to an intermediate code (70I).

116. The method according to any of the preceding embodiments and with the features of embodiment 80, wherein the reference data (50) comprise associated to each of the at least one measurement identifier references (53) a corresponding code specifier (57),

wherein each corresponding code specifier (57) is configured to specify the code (70) that the respective measurement identifier reference (53) corresponds to.

117. The method according to any of the preceding embodiments and with the features of embodiment 108, wherein the reference data (50) are further configured to characterize the plurality of measurements.

118. The method according to the preceding embodiment, wherein the reference data (50) comprise for each measurement at least one, preferably a plurality, of reference characteristic(s) (55).

119. The method according to the preceding embodiment, wherein the at least one reference characteristic (55) comprises at least one of

object, component, substance or specimen to be measured,

unit of measurement,

interval of time over which a measurement is to be made,

object, component, substance or specimen over which the measurement is to be made,

scale type, and

a classification of how the measurement is to be made.

120. The method according to any of the preceding embodiments, wherein configuring the measurement related data (10, 20) according to the second code (70B) comprises

determining for each measurement identifier (3) a replacing measurement identifier reference (53R) and

replacing each measurement identifier (3) with the respective replacing measurement identifier reference (53R).

121. The method according to the preceding embodiment, wherein the replacing measurement identifier reference (53R) is a measurement identifier reference (53) configured according to the second code (70B).

122. The method according to any of the 2 preceding embodiments and with the features of embodiment 80, wherein determining for each measurement identifier (3) a replacing measurement identifier reference (53R)

depends on the respective matching measurement identifier reference (53M) determined for the respective measurement identifier (3).

123. The method according to any of the 3 preceding embodiments and with the features of embodiment 80, wherein determining for each measurement identifier (3) a replacing measurement identifier reference (53R) comprises

determining whether the respective matching measurement identifier reference (53M) determined for the respective measurement identifier (3) is configured according to the second code (70B).

124. The method according to the preceding embodiment and with the features of embodiment 116, wherein the method comprises

utilizing the code specifier (57) corresponding to the matching measurement identifier reference (53M)

to determine whether the respective matching measurement identifier reference (53M) determined for the respective measurement identifier (3) is configured according to the second code (70B).

125. The method according to any of the 2 preceding embodiments, wherein determining for each measurement identifier (3) a replacing measurement identifier reference (53R)

comprises determining the respective replacing measurement identifier reference (53R), for each measurement identifier (3), to be the respective matching measurement identifier reference (53M)

if the matching measurement identifier reference (53M) is determined to be configured according to the second code (70B).

126. The method according to any of the 3 preceding embodiments and with the features of embodiment 110, wherein the method comprises for each detected measurement identifier (3)

finding a measurement identifier reference (53) that is configured according to the second code (70B) and that is associated with the respective matching measurement identifier reference (53M),

if the respective the respective matching measurement identifier reference (53M) is not configured according to the second code (70B).

127. The method according to the preceding embodiment, wherein determining for each measurement identifier (3) a replacing measurement identifier reference (53R)

comprises determining the respective replacing measurement identifier reference (53R), for each measurement identifier (3), to be the measurement identifier reference (53) that is configured according to the second code (70B) and that is associated with the respective matching measurement identifier reference (53M).

128. The method according to any of the preceding embodiments and with the features of embodiment 111,

wherein the first code (70A) is different from the intermediate code (70I) and the second code (70B) is the intermediate code (70I).

That is, the method can comprise configuring the measurement related data (10, 20) from an arbitrary code to the intermediate code (70I), e.g., to a standard code.

129. The method according to any of the preceding embodiments and with the features of embodiment 111,

wherein the first code (70A) is the intermediate code (70I) and the second code (70B) is different from the intermediate code (70I).

That is, the method can comprise configuring the measurement related data (10, 20) from an intermediate code (70I), e.g., a standard code, to an arbitrary code.

130. The method according to any of the preceding embodiments and with the features of embodiment 111,

wherein the first code (70A) is different from the intermediate code (70I) and the second code (70B) is different from the intermediate code (70I).

That is, the method can comprise configuring the measurement related data (10, 20) from a first arbitrary code to a second arbitrary code.

131. The method according to any of the preceding embodiments, wherein the method comprises the data processing system (40) outputting the measurement related data (10, 20) configured according to the second code (70B).

132. The method according to any of the preceding embodiments, wherein the method comprises

a sending node (110) generating the measurement related data (10, 20).

For example, the sending node (110) can be a device/system which can be programmed to generate data configured according to the first code (70A).

133. The method according to the preceding embodiment, wherein the method comprises communicating the measurement related data (10, 20) from the sending node (110) to the data processing system (40).

134. The method according to any of the 2 preceding embodiments, wherein the method comprises communicating the measurement related data (10, 20) from the sending node (110) to the data processing system (40) through an electronic data communication network.

That is, the sending node (110) can be interconnected with the data processing system (40) through an electronic data communication network.

135. The method according to any of the preceding embodiments, wherein the method comprises a receiving node (130) receiving the measurement related data (10, 20) configured according to the second code (70B).

For example, the receiving node (130) can be a device/system which can be programmed to read and/or "understand" data configured according to the second code (70B).

136. The method according to the preceding embodiment, wherein the method comprises communicating the measurement related data (10, 20) from the data processing system (40) to the receiving node (130).

137. The method according to any of the 2 preceding embodiments, wherein the method comprises communicating the measurement related data (10, 20) from the data processing system (40) to the receiving node (130) through an electronic data communication network.

That is, the receiving node (130) can be interconnected with the data processing system (40) through an electronic data communication network.

138. The method according to any of the preceding embodiments, wherein the measurement related data (10, 20) comprise measurement instruction data (10).

139. The method according to the preceding embodiment and with the features of embodiment 132, wherein the sending node (110) is a measurement requesting node.

140. The method according to the preceding embodiment and with the features of embodiment 135, wherein the receiving node (130) is a measurement performing node.

141. The method according to any of the preceding embodiments, wherein the measurement related data (10, 20) comprise measurement result data (20).

142. The method according to the preceding embodiment and with the features of embodiment 132, wherein the sending node (110) is a measurement performing node.

143. The method according to the preceding embodiment and with the features of embodiment 135, wherein the receiving node (130) is a measurement requesting node.

144. The method according to any of the preceding embodiments, wherein the method comprises

a measurement requesting node

generating measurement instruction data (10) configured according to the first code (70A) and

sending the measurement instruction data (10) to the data processing system (40);

the data processing system (40) sending the measurement instruction data (10) configured according to the second code (70B) to a measurement performing node;
the measurement performing node
performing the requested measurement(s),
generating measurement result data (20) configured according to the second code (70B), and
sending the measurement result data (20) to the data processing system (40);
the data processing system configuring the measurement result data (20) according to the first code (70A) and sending the measurement result data (20) to the measurement requesting node.

It will be understood that configuring the measurement related data (10, 20) from the first code (70A) to the second code (70B) is equivalent (in that it comprises similar steps) to configuring the measurement related data (10, 20) from the second code (70B) to the first code (70A).

145. The method according to any of the preceding embodiments, wherein the method is a computer-implemented method.

Below, data processing system embodiments will be discussed. These embodiments are abbreviated by the letter "D" followed by a number. When reference is herein made to data processing system embodiments, these embodiments are meant.

D1. A data processing system (40) comprising
an input unit (401) configured to obtain measurement related data (10, 20) configured according to a first code (70A);
a matching unit (409) configured to
compare the obtained measurement related data (10, 20) with reference data (50) and
to
configure the measurement related data (10, 20) according to a second code (70B) based on the comparison.

D2. The data processing system (40) according to the preceding embodiment, wherein the input unit comprises a network interface.

D3. The data processing system (40) according to any of the 2 preceding embodiments, further comprising an entities recognition unit (403) configured to detect in the obtained measurement related data (10, 20) at least one of

a data element (35), a data portion (30), a measurement identifier (3) and a measurement indicator (2).

D4. The data processing system (40) according to the preceding embodiment, wherein the entities recognition unit (403) is configured to carry out any of the detection steps according to the method embodiments 2 to 42.

D5. The data processing system (40) according to any of the 2 preceding embodiments, wherein the entities recognition unit (403) is configured to obtain the measurement related data (10, 20) from the input unit (401).

D6. The data processing system (40) according to any of the 3 preceding embodiments, wherein the data processing system (40) comprises an online pre-processing unit (405) configured to pre-process the at least one detected measurement identifier (3).

D7. The data processing system (40) according to the preceding embodiment, wherein the entities recognition unit (403) is configured to detect at least one measurement identifier (3) and provide the measurement identifier (3) to the online pre-processing unit (405).

D8. The data processing system (40) according to any of the 2 preceding embodiments, wherein the online pre-processing unit (405) is configured to carry out any of the pre-processing steps according to the method embodiments 43 to 79.

D9. The data processing system (40) according to any of the 8 preceding embodiments, wherein the matching unit (409) is configured to carry out any of the steps according to the method embodiments 80 to 107 and 120 to 131.

D10. The data processing system (40) according to any of the 8 preceding embodiments, wherein the data processing system (40) comprises a similarity metric calculating unit (407) configured to receive two inputs and to calculate a similarity metric between the first input and the second input.

D11. The data processing system (40) according to the 2 preceding embodiments, wherein the matching unit (409) is configured to utilize the similarity metric calculating unit (407) to carry out any of the steps according to the method embodiments 96 and 105 to 107.

D12. The data processing system (40) according to any of the 11 preceding embodiments, wherein the data processing system (40) comprises an offline pre-processing unit (407) configured to pre-process the reference data (50).

D13. The data processing system (40) according to the preceding embodiment, wherein the reference data (50) comprise a plurality of identifier references (53) and
wherein the offline pre-processing unit (413) is configured to pre-process the identifier references (53).

D14. The data processing system (40) according to the preceding embodiment and with the features of embodiment D6, wherein the offline pre-processing unit (413) is configured to carry out the same pre-processing steps on the identifier references (53) as the online pre-processing unit (413) does on the measurement identifier(s) (3).

D15. The data processing system (40) according to any of the 14 preceding embodiments, wherein the reference data (50) are configured according to embodiments 108 to 119.

D16. The data processing system (40) according to any of the 15 preceding embodiments, wherein the data processing system (40) comprises an output unit (415) configured to output the measurement related data (10, 20) configured according to the second code (70B).

D17. The data processing system (40) according to the preceding embodiment, wherein the output unit (415) comprises at least one of a display, a printer, a fax, and a network card.

D18. The data processing system (40) according to any of the 17 preceding embodiments, wherein the data processing system (40) comprises a learning unit (411) configured to extend the reference data (50).

D19. The data processing system (40) according to any of the 18 preceding embodiments, wherein the data processing system is configured to carry out the method according to any of the preceding method embodiments.

Below further method embodiments will be discussed.

146. The method according to any of the preceding method embodiments, wherein the data processing system (40) is configured according to any of the data processing system embodiments.

Below, communication system embodiments will be discussed. These embodiments are abbreviated by the letter "S" followed by a number. When reference is herein made to communication system embodiments, these embodiments are meant.

S1. A communication system comprising a data processing system (40) wherein the data processing system (40) is configured to
 obtain measurement related data (10, 20) configured according to a first code (70A);
 compare the obtained measurement related data (10, 20) with reference data (50);
 configure the measurement related data (10, 20) according to a second code (70B)
based on the comparison.

S2. The communication system according to the preceding system embodiment, wherein the communication system comprises a memory component configured to store the reference data (50).

S3. The communication system according to the preceding embodiment, wherein the data processing system (40) is configured to access the memory component.

S4. The communication system according to any of the 2 preceding embodiments, wherein the memory component is integrated into the data processing system (40).

S5. The communication system according to any of the preceding communication system embodiments, further comprising a sending node (110) configured to generate the measurement related data (10, 20) according to the first code (70A).

S6. The communication system according to the preceding embodiment, wherein the sending node (110) is configured to electronically communicate with the data processing system (40).

S7. The communication system according to any of the preceding communication system embodiments, further comprising a receiving node (130) configured to receive the measurement related data (10, 20) configured according to the second code (70B).

S8. The communication system according to the preceding embodiment, wherein the receiving node (130) is configured to electronically communicate with the data processing system (40).

S9. The communication system according to any of the preceding communication system embodiments, wherein the communication system is configured to carry out the method according to any of the preceding method embodiments.

S10. The communication system according to any of the preceding communication system embodiments and with the features of embodiments S5,

wherein the data processing system (40) is configured to carry out the method according to any of the preceding method embodiments

to configure the measurement related data (10, 20) generated by the sending node (110) according to the second code (70B).

S11. The communication system according to the preceding embodiment, wherein the second code (70B) is an intermediate code (70I).

S12. The communication system according to any of the preceding communication system embodiments and with the features of embodiments S5 and S7,

wherein the data processing system (40) is configured to carry out the method according to any of the preceding method embodiments

to facilitate the communication between the sending node (110) and the receiving node (130).

S13. The communication system according to any of the preceding communication system embodiments, wherein the data processing system (40) is configured according to any of the data processing system embodiments.

Below further numbered embodiments are discussed.

C1. A computer program product comprising instructions, which when the program is executed by a computer can cause the computer to carry out the method according to any of the preceding method embodiments.

C2. A computer program product comprising instructions, which when the program is executed by the data processing system (40) can cause the data processing system (40) to carry out the method according to any of the preceding method embodiments.

R1. A computer-readable storage medium comprising instructions, which when the instructions are executed by a computer can cause the computer to carry out the method according to any of the preceding method embodiments.

R2. A computer-readable storage medium comprising instructions, which when the instructions are executed by the data processing system (40) can cause the data processing system (40) to carry out the method according to any of the preceding method embodiments.

U1. Use of the method and/or data processing system (40) and/or communication system according to any of the preceding method and/or data processing system (40) and/or communication system embodiments for configuring measurement related data (10, 20) from a first code (70A) to a second code (70B).

U2. Use of the method and/or data processing system (40) and/or communication system according to any of the preceding method and/or data processing system (40) and/or communication system embodiments for configuring measurement related data (10, 20) from a first code (70A) to a second code (70B), wherein the second code (70B) is an intermediate code (70I).

U3. Use of the method and/or data processing system (40) and/or communication system according to any of the preceding method and/or data processing system (40) and/or communication system embodiments for facilitating a communication between

a sending node (110) configured to generate measurement related data (10, 20) according to a first code (70A) and

a receiving node (130) configured to receive measurement related data (10, 20) according to a second code (70B).

U4. Use of the method and/or data processing system (40) and/or communication system according to any of the preceding method and/or data processing system (40) and/or communication system embodiments for facilitating a transmission of measurement related data (10, 20) from a sending node (110) configured to generate measurement related data (10, 20) according to a first code (70A) to a receiving node (130) configured to receive measurement related data (10, 20) according to a second code (70B).

The present invention will now be described with reference to the accompanying drawings which illustrate embodiments of the invention. These embodiments should only exemplify, but not limit, the present invention.

- | | |
|-----------------|--|
| Fig. 1 | illustrates a communication system according to an embodiment of the present invention; |
| Figs. 2a and 2b | illustrate measurement related data; |
| Fig. 3 | illustrates reference data that can be utilized while processing measurement related data; |
| Figs. 4a to 4g | illustrate a method of processing measurement related data; |
| Fig. 5 | graphically illustrates an iterative comparison of the measurement related data with the reference data; |
| Fig. 6 | graphically illustrates the method of processing measurement related data; |
| Fig. 7 | illustrates measurement related data after processing; |
| Figs. 8a to 8d | depict examples of reference data, measurement related data and an output of the method; |
| Fig. 9a | depicts an example of measurement related data; |
| Fig. 9b | illustrates an exemplary output after processing the measurement related data of Fig. 9b; |
| Fig. 10 | illustrates units that can be comprised by a data processing system. |

In the following, exemplary embodiments of the invention will be described, referring to the figures. These examples are provided to give further understanding of the invention, without limiting its scope.

In the following description, a series of features and/or steps are described. The skilled person will appreciate that unless explicitly required and/or unless required by the context, the order of features and steps is not critical for the resulting configuration and its effect. Further, it will be apparent to the skilled person that irrespective of the order of features and steps, the presence or absence of time delay between steps can be present between some or all of the described steps.

It is noted that not all the drawings carry all the reference signs. Instead, in some of the drawings, some of the reference signs have been omitted for sake of brevity and simplicity of illustration. Embodiments of the present invention will now be described with reference to the accompanying drawings.

Fig. 1 depicts a communication system comprising a data processing system 40 configured to process measurement related data 10, 20.

The data processing system 40 may comprise one or more processing units configured to carry out computer instructions of a program (i.e. machine readable and executable instructions). The processing unit(s) can be singular or plural. For example, the data processing system 40 may comprise at least one of central processing unit (CPU), graphical processing unit (processing unit) GPU, digital signal processor (DSP), accelerated processing unit (APU), application specific integrated circuit (ASIC), application specific instruction set processor (ASIP), field programmable gate array (FPGA), artificial intelligence (AI) accelerator and tensor core, each of which can be in the singular or plural.

The data processing system 40 may comprise one or more memory component(s), such as, main memory (e.g. RAM), cache memory (e.g. SRAM) and/or secondary memory (e.g. HDD, SDD). The data processing system 40 may comprise volatile and/or non-volatile memory, such, a synchronous RAM (SRAM), dynamic RAM (DRAM), synchronous dynamic RAM (SDRAM), Flash Memory, magnetic (MRAM), ferroelectric RAM (FRAM) or parameter RAM (P-RAM).

The data processing system 40 may comprise one or more internal communication interface(s) or component(s) (e.g. busses) configured to facilitate electronic data exchange between the

components of the data processing system 40, such as, the communication between the memory components and the processing components.

The data processing system 40 may comprise one or more external communication interface(s) or component(s), configured to facilitate electronic data exchange between the data processing system 40 and devices or networks external to the data processing system 40. For example, the external communication component can comprise at least one network interface card which can be configured to connect the data processing system 40 to a (wired and/or wireless) network, such as, to the Internet and/or to a cellular network. The external communication component can be configured to transfer electronic data using a standardized communication protocol (e.g. TCP/IP protocol). The data processing system 40 may be a centralized or distributed computing system. For example, the data processing system 40 may be configured for providing cloud computing functionalities. Alternatively or additionally, the data processing system 40 may be provided locally, e.g., in a laboratory.

The data processing system 40 can comprise user interfaces, e.g., output user interface, such as, displays, screens, monitors, speakers and input user interface, such as, a keyboard, trackpad, mouse, touchscreen and joystick.

The data processing system 40 can be configured to carry out instructions of one or more computer program(s). The data processing system 40 can be a system-on-chip comprising one or more processing component(s), memory component(s), internal data communication component(s) and external data communication component(s). The data processing system 40 can be a personal computer, a laptop, a pocket computer, a smartphone or a tablet computer. The data processing system 40 can be a server, a server system, a portion of a cloud computing system or a system emulating a server, such as, a server system with an appropriate software for running a virtual machine. The data processing system 40 can be a processing component or a system-on-chip that can be interfaced with a personal computer, a laptop, a pocket computer, a smartphone, a tablet computer and/or user interface (such as the upper-mentioned user interfaces). The latter can facilitate extending the functionalities of existing systems with the functionalities of the data processing system 40 discussed below. For example, the data processing system 40 can be easily integrated into the existing systems and/or devices of a laboratory, such as, of a clinical laboratory.

The data processing system 40 can be configured to obtain or to have access to reference data 50 (discussed in detail with reference to Fig. 3). For example, the reference data 50 can be stored in a memory device (not shown) that can be accessed by or integrated into or comprised by the data processing system 40. In addition, the data processing system 40 can

be configured to receive (i.e. obtain) measurement related data 10, 20 configured according to a first code 70A and to compare the obtained measurement related data 10, 20 to the reference data 50. Based on the comparison, the data processing system 40 can configure the measurement related data 10, 20 according to a second code 70B.

The measurement related data 10, 20 may relate to one or more measurements. The measurement(s), as used throughout this document, can refer to measuring in the classical sense and to chemical, bio-chemical or biologic analysis of an object or component, which can generate information about at least one physical, chemical, biological or bio-medical feature of the object or component. The object or component can for example be a sample. The samples may be samples originating from the human body. For example, the samples may comprise samples of a bodily fluid, such as blood, urine or saliva. The samples may, for example, also comprise samples of at least one tissue from the human body, such as from a biopsy.

However, depending on the use of the present invention, the measurement can also refer to measurements performed in other technological environments, e.g., size measurements of a product (e.g. in quality control), temperature or pressure measurements in an industrial environment (e.g. in a controlled industrial environment), amount of a substance in an environment, in the air or in a mixture (e.g. toxicity measurements), measurement of networking parameters (e.g. packet length, packet delay, bandwidth) etc.

The measuring can be direct, such as, measuring a temperature by means of a sensor device configured for sensing the temperature. It can also be indirect, i.e., determining data based on sensed data. For example, an acceleration can be indirectly measured by measuring a corresponding velocity and generating a derivate with respect to the time, or by an adapted sensor device that senses a force of an object with known inertia against the accelerated object. In another example, a concentration of a substance A in a mixture B is measured by determining the amount of substance A in a known or determined amount of substance B. This can be the case, e.g., when measuring a concentration of cholesterol in blood of a user.

The measurement related data 10, 20 can be generated, e.g., by a sending node 110 and can be received, e.g., by a receiving node 130.

The measurement related data 10, 20 can comprise, inter alia, data configured to specify a measurement. The measurement related data 10, 20 may comprise, e.g., a name of a measurement, a number, a value, a range, a range specifier, a unit, a component or sample

name on which a measurement is performed, a method used to perform the measurement and a duration of the measurement.

The measurement related data 10, 20 can comprise measurement instruction data 10, which can also be referred to as measurement requesting data 10. The measurement instruction data 10 can comprise instructions for performing a measurement. More particularly, the measurement instruction data 10 can comprise at least one request for a measurement. Thus, the measurement instruction data 10 can comprise at least the name of a measurement requested to be performed. However, the measurement instruction data 10 can also comprise further data characterising or further specifying the requested measurement, e.g., the unit of the measurement, a component on which the measurement is to be performed, the duration of the measurement and a method for performing the measurement.

For example, the measurement instruction data 10 can be generated by a measurement requesting device or system (e.g. a device or system used by a physician or by a patient). Moreover, the measurement instruction data 10 can comprise instructions for instructing a measurement performing device or system (e.g. the LIS of a laboratory and/or a device configured to perform the measurement) to perform a measurement. In yet another example, a control system may generate measurement instruction data 10 for instructing a sensor device to perform a measurement. The instructions that can be comprised by the measurement instruction data 10 can comprise computer instructions (i.e. machine-readable instructions). Alternatively or additionally, the instructions that can be comprised by the measurement instruction data 10 can be converted (e.g. compiled) into computer instructions (i.e. machine-readable instructions). Alternatively or additionally, the instructions that can be comprised by the measurement instruction data 10 can comprise or can be converted into human-readable instructions.

Thus, typically, measurement instruction data 10 can be generated by a node (e.g. sending node 110) that requests a measurement and can be transmitted to another node (e.g. receiving node 130) that can perform the measurement. The measurement instruction data 10 can be communicated from the measurement requesting node (e.g. sending node 110) to the measurement performing node (e.g. receiving node 130) via the data processing system 40. This can allow the data processing system 40 to increase the interoperability between the two nodes (e.g. between the sending node 110 and the receiving node 130).

The term node, as used in this document, can be analogous to a network node in an electronic data communication network. Moreover, the nodes 110, 130 can comprise hardware and/or software components configured to process and/or store and/or communicate electronic data.

For example, the nodes 110, 130 can comprise a workstation, computer, tablet, mobile phone, server, a laboratory information system and/or a hospital information system.

For example, the measurement requesting node (e.g. sending node 110) can be configured to generate (and in general to store, process and communicate) data according to a respective code 70, e.g., according to the first code 70A. On the other hand, the measurement performing node (e.g. receiving node 130) can be configured to generate (and in general to store, process and communicate) data according to a respective code 70, e.g., according to the second code 70B. While otherwise, the communication between the measurement requesting node and the measurement performing node would be error-prone or even impossible (due to lack of code compatibility), the data processing system 40 can alleviate this issue and can thus facilitate a successful communication between the two nodes.

For example, the measurement requesting node can generate measurement instruction data 10 that can comprise the instruction "Perform HDL Chol. in Sample X". On the other hand, the measurement performing node can use a different code for referring to measurements and can refer to the said measurement with, e.g., the code "14646-4". As such, the measurement performing node would fail to understand (e.g. decode or compile) or properly process (e.g. execute) the instruction "Perform HDL Chol. in Sample X". For example, the LIS of the measurement performing node would not be able to process said instruction. However, the data processing system 40 may alleviate this issue, by receiving the instruction "Perform HDL Chol. in Sample X" from the measurement requesting node and by instructing the measurement performing node to perform "14646-4 on X". Thus, a successful communication between the two nodes can be realized. In the above example, X can be an identification code configured to identify a component, object, or sample to be measured.

The measurement related data 10, 20 can comprise measurement result data 20. The measurement result data 20 can comprise results obtained by performing a measurement. More particularly, the measurement result data 20 can comprise at least the name of a measurement that is performed, a number, a value, a range, a range specifier, and a unit indicating the obtained measurement result(s). The measurement result data 20 can also comprise further data characterising or further specifying the performed measurement, e.g., a component name on which the measurement is performed, the duration of the measurement and a method used to perform the measurement.

For example, the measurement result data 20 can be generated by a measurement performing device or system (e.g. a laboratory and/or device performing a measurement). Moreover, the measurement result data 20 can comprise results obtained by performing a measurement,

e.g., a measurement requested by a measurement requesting device or system (e.g. a device or system used by a physician or by a patient). In yet another example, a sensor device may generate the measurement result data 20 after performing a measurement, e.g., requested by a control system. The measurement result data 20 can comprise machine readable data indicating the measurement and/or human readable data.

Thus, typically, the measurement result data 20 can be generated by a node that performs a measurement and can be transmitted to a node that requested the measurement. The measurement result data 20 can be communicated from the measurement performing node (in this example acting as the sending node 110) to the measurement requesting node (in this example acting as the receiving node 130) via the data processing system 40. This can allow the data processing system 40 to increase the interoperability between the two nodes 110, 130.

For example, the measurement performing node can be configured to generate (and in general to store, process and communicate) data according to a respective code 70, e.g., according to the first code 70A. On the other hand, the measurement requesting node can be configured to generate (and in general to store, process and communicate) data according to a respective code 70, e.g., according to the second code 70B. Whilst otherwise the communication between the measurement performing node and the measurement requesting node would be error-prone or even impossible (due to lack of code compatibility), the data processing system 40 can alleviate this issue and can thus facilitate a successful communication between the two nodes.

For example, the measurement performing node can generate measurement result data 20 that can comprise the measurement result (in human and/or machine-readable data) "HDL Chol.: 45 mg/dL". On the other hand, the measurement requesting/receiving node can use a different code for referring to measurements and can refer to the said measurement with the code "14646-4 [mmol/L]". As such, the measurement requesting/receiving node would fail to understand or properly process the result "HDL Chol.: 45 mg/dL". For example, the laboratory information system of the measurement requesting/receiving node would not be able to process said result. However, the data processing system 40 may alleviate this issue, by receiving the result "HDL Chol.: 45 mg/dL" from the measurement performing node and by providing to the measurement requesting/receiving node the result "14646-4: 1.1637 mmol/L". Thus, the utilization of the data processing system 40 can facilitate a successful communication between the two nodes.

The measurement related data 10, 20 can be electronic data. This can allow the data processing unit 40 to process the measurement related data 10, 20. In addition, this can allow the measurement related data 10, 20 to be transmitted using electronic data transmission systems.

Throughout this document, the term measurement related data 10, 20 can be used to jointly refer to the measurement instruction data 10 and to the measurement results data 20.

The measurement requesting device or system can also be referred to as a measurement requesting node. The measurement performing device or system can also be referred to as a measurement performing node. Moreover, depending on the scenario the measurement requesting node may be a sending node 110 or a receiving node 130. Similarly, the measurement performing node may be a sending node 110 or a receiving node 130. Typically, when the measurement related data 10, 20 comprise measurement instruction data 10, the measurement requesting node is a sending node 110 and the measurement performing node is a receiving node. Alternatively, when the measurement related data 10, 20 comprise measurement result data 20, the measurement performing node is a sending node 110 and the measurement instructing node is the receiving node 130. It will be understood, that in any case, the receiving node 130 may be another node in addition to the measurement performing node and the measurement instructing node. For example, the measurement result data 20 may be provided to a laboratory requesting the measurement and/or to a patient to which the measurement concerns. Although for brevity the present invention is described with one sending node 110 and with one receiving node 130, it will be understood that the present invention can similarly be utilized by multiple sending nodes 110 and/or receiving nodes 130. The term nodes can be used herein to refer to network nodes and more particularly to network nodes in an electronic data communication network.

Figs. 2a and 2b illustrate the measurement related data 10, 20. More particularly, Figs. 2a and 2b illustrate a structure of the measurement related data 10, 20. That is, the measurement related data 10, 20 can be configured as a data structure (or object) that can comprise further data structures (or objects). For example, the measurement related data 10, 20 can be modular data structures or can be configured (e.g. by the data processing device 40, see Fig. 1) as modular data structures. This can facilitate the storing and processing of the measurement related data 10, 20. Moreover, it can facilitate the detection of measurement identifiers 3 (described further below).

The measurement related data 10, 20 may comprise, inter alia, data indicating one or more measurement(s), e.g., name(s), number(s), value(s), range specifier(s), unit(s), component

name(s), method(s) and duration(s). Corresponding to each of the said elements, the measurement related data 10, 20 can comprise one or more data element(s) 35. That is, at least some of the data elements 35 may indicate one or more measurement(s), e.g., name(s), number(s), value(s), range specifier(s), unit(s), component name(s), method(s) and duration(s).

Each data element 35 may comprise or may consist of at least one character, such as, ASCII characters. This can facilitate configuring the data elements 35 as machine readable data and as human readable data.

Moreover, each data element 35 can be detectable. For example, the measurement related data 10, 20 can comprise a plurality of element delimiters 37 which can be configured to indicate a boundary of a data element 35. For example, the data elements 35 can be words and the element delimiters 37 can be space characters and/or punctuation marks.

Alternatively or additionally, the measurement related data 10, 20 may comprise indices or keys which can be used to differentiate the data elements 35 from each other. For example, the measurement related data 10, 20 can be an ordered list of data elements 35, wherein each data element 35 can be associated with a respective index indicating the position of the data element in the ordered list. In other words, each data element 35 can be stored in a respective location in a memory, which location can be determined based on the respective index of the data element 35. Alternatively, other data structures may be used to store the measurement related data 10, 20, e.g., linked lists, arrays, vectors, matrices, etc.

The measurement related data 10, 20 may further comprise one or more data portion(s) 30. The data portions 30 can comprise data elements 35. Typically, the data portions 30 can be configured to organize the data elements 35. This can facilitate the intelligibility of the measurement related data 10, 20. Additionally, it can facilitate detecting at least one measurement identifier 3 (discussed further below). For example, the data portions 30 may correspond to respective blocks of data (e.g. lines) of the measurement related data 10, 20. Moreover, the data portions 30 may comprise a different number of data elements 35. Some data portions 30 may be empty, i.e., may not comprise any data element 35.

Each data portion 30 can be detectable. For example, the measurement related data 10, 20 can comprise element delimiter(s) 33 configured to indicate a boundary of a data portion 30. For example, the data portion 30 can correspond to a line of the measurement related data 10, 20 and the element delimiter 33 can be a new line character.

The measurement related data 10, 20 can be configured, e.g., in a tabular format (e.g. see Fig. 8c). Each row of the table can be a data portion 30 and each cell of the table can be occupied by a data element 35. Some cells of the table can be unoccupied.

In some embodiments, each data portion 30 can be configured to correspond to a respective measurement. That is, in some embodiments, the data comprised in a data portion 30 correspond to (e.g. identify and specify) one measurement. However, some data portions 30 may not correspond to or indicate a measurement. For example, some data portions 30 may comprise auxiliary data, such as, a name of a node that can generate/receive the measurement related data 10, 20, an address and a phone number.

As illustrated in Fig. 2b, the measurement related data 10, 20 may comprise at least one measurement identifier 3. The measurement identifier 3 can be configured to identify a measurement. For example, the measurement identifier 3 may comprise the name of a measurement or an ID sequence for identifying a measurement. Each measurement identifier 3 may comprise one or more data elements 35. In other words, one data element 35 or a sequence of data elements 35 may correspond to or may form a measurement identifier 3.

In addition, the measurement related data 10, 20 may comprise one or more measurement indicators 2. This can typically be the case when the measurement related data 10, 20 comprise measurement result data 20. The measurement indicators 2 may typically comprise data that can further specify and/or characterise a measurement, such as, numbers, values, ranges, and units. For example, in the case of measurement result data 20, each measurement identifier 3 can be associated with numbers, ranges and units for indicating the result of the measurement. In the case of measurement instruction data 10, each measurement identifier 3 can be associated with range specifiers or units that the measurement results are requested to be provided. As such, the presence of numbers, ranges, range specifiers and units can typically indicate the presence of a measurement identifier 3. Again, this is based in the rationale that a measurement can typically be provided in the measurement related data 10, 20 with a name and with further specifying data such as numbers, values, unit, ranges, component names and method types. The measurement indicators 2 can comprise one or more data elements 35. In other words, one data element 35 or a sequence of data elements 35 may correspond to or may form a measurement indicator 2.

That is, some of the data elements 35 may indicate measurement data. Typically, said measurement data elements 35 can be measurement names or IDs, numbers, values, units, ranges, range specifiers, method names or IDs and component names or IDs. As such, typically said data elements 35 can be part of measurement identifiers 3 or measurement

indicators 2. However, some of the data elements 35' may not convey any information that can relate to a measurement. Typically, such data elements 35' may be, e.g., stop words, punctuation marks, a name of a node that can generate/receive the measurement related data 10, 20, an address and a phone number.

In some embodiments, the data processing system 40 (see Fig. 1) can be configured to detect the measurement identifiers 3 and the measurement indicators 2 by detecting and classifying the data elements 35 correspondingly. That is, the data processing system 40 can be configured to detect data elements 35 and to determine for each data element 35 whether it is part of a measurement identifier 3 or measurement indicator 2 or neither. In the latter, the data processing system 40 can be configured to determine whether a data element 35 is an irrelevant data element 35' that does not convey any information that can relate to a measurement.

Fig. 3 depicts a schematic illustration of the reference data 50.

As discussed with reference to Fig. 1, the data processing system 40 can compare the obtained measurement related 10, 20, configured according to a first code 70A, with the reference data 50. Based on the comparison, the data processing system 40 can configure the measurement related data 10, 20 according to a second code 70B. That is, the reference data 50 can facilitate configuring the measurement related data 10, 20 from a first code 70A to a second code 70B.

The reference data 50 can be configured to relate to a plurality of measurements. The plurality of measurements may typically correspond to a field of technology. For example, the reference data 50 may relate to a plurality of bio-medical measurements (i.e. medical laboratory observations). Such reference data 50 may be used to process measurement related data 10, 20 relating to, e.g., bio-medical measurements.

More particularly, the reference data 50 can be configured to identify a plurality of measurements. That is, the reference data 50 can comprise data that can be used to refer to a measurement. More particularly, the reference data 50 can comprise a plurality of measurement identifier references 53, wherein each measurement identifier reference 53 can be used (e.g. by a node, a laboratory, a LIS, a HIS, a physician, a hospital) to refer to a measurement. For example, the reference data 50 can comprise a plurality of biomarkers 53 that can facilitate identifying medical laboratory observations.

In a preferred embodiment, the reference data 50 can comprise for each measurement a plurality of measurement identifier references 53. For example, for each measurement, the

reference data 50 can comprise a plurality of names, synonyms, abbreviations and/or code names that different nodes (e.g. laboratories) can use to refer to the measurement. In other words, the reference data 50 can comprise, for each measurement, a plurality of measurement identifier references 53 each corresponding to at least one respective code 70. Generally, having a large number of measurement identifier references 53 for each measurement can be advantageous, as it can facilitate detecting measurement identifiers 3 in measurement related data 10, 20.

Furthermore, at least one of the measurement identifier references 53 that correspond to a measurement can be a measurement identifier reference 53 configured to uniquely and/or unambiguously identify the measurement (e.g. see first column of the table in Fig. 8a). Said measurement identifier reference 53 can comprise, e.g., standard (code) names or universal (code) names. That is, for each measurement, the reference data 50 can comprise at least one measurement identifier reference 53 which is configured according to an internal code 70I (which can also be referred to as an intermediate code 70I). The internal code 70I can be configured to unambiguously or uniquely refer to measurements. Typically, the internal code 70I can be a standard code. For example, the reference data 50 can comprise a plurality of measurement identifier references configured according to the Logical Observation Identifiers Names and Codes (LOINC). LOINC is a database and universal standard for identifying medical laboratory observations. It uses universal code names and identifiers to medical terminology (e.g. see Fig. 8a).

Thus, the reference data 50 can comprise for each measurement a plurality of measurement identifier references 53, at least one of which can be a standard identifier and the rest can be non-standard, non-universal and/or node-specific identifiers for a measurement. This can be advantageous as typically nodes (e.g., laboratories) use internal or non-universal identifiers to refer to measurements. As such, e.g., a laboratory may generate measurement related data 10, 20 comprising measurement identifiers 3 that are used only locally or internally by the laboratory. As such, the reference data 50 can facilitate unambiguously identifying the measurement(s) that the measurement related data 10, 20 refer to. For example, the reference data can facilitate matching (i.e. mapping) the (laboratory-specific) measurement identifiers 3 to a measurement identifier reference 53 which can be standard or universal (i.e. configured according to the internal code 70I). For example, the reference data 50 can facilitate matching a laboratory-specific biomarker 3 to a LOINC code name 53 referring to the same measurement. Similarly, the reference data 50 can facilitate matching a laboratory-specific biomarker 3 used by a first laboratory with another laboratory specific biomarker 53 used by a second laboratory and comprise in the reference data 50.

In addition to identifying measurements, the reference data 50 can be configured to characterize measurements. More particularly, for each measurement, the reference data 50 can comprise at least one reference characteristic 55 configured to characterise the measurement. The at least one reference characteristic 55 can comprise or can be a unit, a list of units, a component to be measurement, a component or sample on which the measurement is to be performed, an interval or length of a measurement, a scale type, a method of how to perform a measurement or any combination thereof. That is, the reference data 50 can comprise for each measurement at least one reference characteristic 55 configured to specify how to provide the results of the measurement.

In addition, the reference data 50 can comprise code specifiers 57 which can be configured to specify the code 70 that the at least one measurement identifier reference 53 belongs to. This can be particularly advantageous for configuring the obtained measurement related data 10, 20 according to a particular code 70. For example, the code specifier 57 can indicate, for each measurement identifier reference 53, whether it corresponds to the internal (e.g. standard) code 70I or to a non-universal code 70. In some embodiments, the code specifier 57 can indicate, for each measurement identifier reference 53, a node or a list of nodes that use the measurement identifier reference 53. The latter can facilitate determining the corresponding measurement identifier reference 53 that can be used to replace a respective measurement identifier 3, such that, the measurement related data 10, 20 can be configured according to the second code 70B.

Moreover, the reference data 50 can comprise a data cleaning database 510, an equivalent data elements database 540 and auxiliary data 560.

The data cleaning database 510 can facilitate pre-processing the measurement identifiers 3 and more particularly a data cleaning step (see Fig. 4f). The data cleaning database 510 can facilitate detecting in the measurement identifiers 3 irrelevant data elements 35' which do not facilitate the identification of a measurement (e.g., stop-words and punctuation marks). Thus, the data cleaning database 510 can comprise a plurality of (possible) irrelevant data elements 35' (e.g. stop words). Hence, the data cleaning database 510 can be used to determine whether a data element 35 is irrelevant or not (e.g. based on whether the data element can be found on the data cleaning database 510).

The equivalent data elements database 540 can facilitate pre-processing the measurement identifiers 3 and more particularly a data element replacement step (see Fig. 4f). The equivalent data elements database 540 can comprise a plurality of data elements, each associated with at least one equivalent data element 35E. This facilitate replacing a data

element 35 with at least one of the associated equivalent data elements 35E. For example, the equivalent data elements database 540 can comprise a dictionary of synonyms (i.e. synonyms dictionary).

The auxiliary data 560 can comprise thresholds and/or ranges which can be used during the comparison of the measurement related data 10, 20 with the reference data 50.

Figs. 4a to 4g illustrate a method of processing measurement related data 10, 20.

Fig. 4a depicts general steps of the method. In a step S1 the method can comprise obtaining measurement related data 10, 20 configured according to a first code 70A. As discussed, the measurement related data 10, 20 may comprise measurement instruction data 10 and/or measurement result data 20. In a step S3, the method can comprise comparing the obtained measurement related data 10, 20 configured according to the first code 70A with reference data 50. The data processing system 40 can perform the comparison. In a step S4, the method can comprise configuring the measurement related data 10, 20 according to a second code 70B, based on the comparison.

In some embodiments, the method can comprise step S2 (Fig. 4b), detecting at least one measurement identifier 3 in the obtained measurement related data 10, 20. Typically step S2 can be performed prior to step S3. In such embodiments, in step S3 the method can comprise comparing the detected measurement identifier 3 with the reference data 50 (step S31).

Fig. 4c depicts steps of detecting at least one measurement identifier 3 in the obtained measurement related data 10, 20 according to an embodiment. More particularly, Fig. 4c illustrates a direct detection of the measurement identifiers 3 in the measurement related data 10, 20. In such embodiments, the step of detecting at least one measurement identifier 3 (i.e. step S2) can comprise detecting a plurality of data elements 35 comprised by the measurement related data 10, 20 (i.e. step S21). Further, detecting at least one measurement identifier 3 (i.e. step S2) can comprise determining for each data element 35 whether it corresponds to a measurement identifier 3 (i.e. step S22). For example, the data processing system 40 can detect a plurality of data elements 35 and based on the reference data 50 can determine for each data element whether it is measurement identifier 3 or part of a measurement identifier 3. For example, the data processing system 40 can compare each data element 35 or a sequence of data elements 35 with the reference data 50 to find a matching measurement identifier references 53M (see Fig. 6). More particularly, the data processing system 40 can compare each data element 35 or a sequence of data elements 35 with the plurality of identifier references 53 comprised by the reference data 50. If a high similarity is

determined (obtained/resulted) from the comparison, that is, if at least one matching identifier reference 53M can be determined based on the comparison, then the data processing system 40 can determine that the data element 35 or the sequence of data element 35 can be (part) of a measurement identifier 3.

However, this may not be very efficient as it can require a larger number of comparisons (i.e. each data element 35 may need to be compared with the reference data to detect measurement identifier references 3). This can be particularly disadvantageous if the number of identifier references 53 comprised by the reference data 50 is large, hence, making a comparison with the reference data 50 a computationally expensive operation. On the other hand, directly detecting the measurement identifiers 3 by comparing the data elements 35 with reference data (i.e. steps S21 and S22) can be advantageous as the detecting the measurement identifier 3 and matching them with matching identifier references 53M can be performed simultaneously. That is, when a high similarity between a data element 35 or a sequence of data elements 35 with an identifier reference 53 is found, not only the measurement identifier 3 is detected but, in addition, the respective identifier reference 53 can be determined to be the matching identifier reference 53M to the detected measurement identifier 3).

Fig. 4d depicts steps of detecting at least one measurement identifier 3 according to an alternative embodiment to the one depicted in Fig. 4c. More particularly, Fig. 4d illustrates an indirect detection of the measurement identifiers 3. Again, in a step S2, detecting at least one measurement identifier 3 (i.e. step S2) can comprise step S21, detecting a plurality of data elements 35 comprised by the measurement related data 10, 20. The detection of the data elements 35 can be facilitated by the detection of the element delimiters 37 (see Fig. 2a). In a step S23, for each data element 35 it can be determined whether it corresponds to a measurement indicator 2. For example, it can be determined whether the data element 35 is at least one of a number, a value, a unit, a range, a name of a measurement method and a name of a measured component. Typically, a measurement can be provided with a number, value, unit, range, method and component. Thus, their presence typically can indicate the presence of a measurement identifier 3. Thus, in a step S24 the method can comprise detecting at least one measurement identifier 3 based on the detection of at least one measurement indicator 2.

Generally, the indirect detection of the measurement identifiers (Fig. 4d) can be more efficient than the direct detection of the measurement identifiers (Fig. 4c). This is due to the fact that it can be less computationally complex to determine whether a data element is, e.g., a

number, value, unit or range, then to determine that the data element matches to measurement identifier reference 53. For example, it can require less computations to determine whether a data element 35 is a number, a unit or a range specifier than to determine that a data element 35 matches to a measurement identifier reference 53. This can be particularly the case if the reference data 50 comprise a large database of identifier references 53.

Fig. 4e depicts an example of the indirect detection of the measurement identifiers discussed in Fig. 4d. In a step S25, the method can comprise detecting a plurality of data portions 30 (see Fig. 2) comprised by the measurement related data 10, 20. The detection of the data portions 30 can be facilitated by detecting the portion delimiters 33 (see Fig. 2a). In a step S26, for each data portion 30, it can be determined whether the data portion 30 comprises a measurement indicator 2. For example, for each data element 35 comprised by the data portion 30 it can be determined whether it corresponds to a measurement indicator 2 (e.g. whether the data element 35 is a number, unit or range specifier). In a step S27, upon detection of a measurement indicator 2, it can be determined that the rest of the data comprised in the corresponding data portion 30 comprises a measurement identifier 3. For example, if in a data portion 30 (e.g. a line) of the measurement related data 30 at least one number or unit is detected, then it can be determined that the data portion 30 comprises a measurement identifier 3. Moreover, it can be determined that all the data elements 35 of the data portion 30 that do not correspond to a measurement indicator 2, correspond to a measurement identifier 3.

Put simply, in some embodiments, the measurement identifiers 3 can be detected by determining whether the data portions 30 of the measurement related data 10, 20 comprise a measurement indicator 2, e.g., number, unit or range specifier. If at least one measurement indicator 2 is detected in a data portion 30, the rest of the data portion 30 can be hypothesized to be the measurement identifier 3.

However, in some instances the rest of the data portion 30 that is hypothesized to be the measurement identifier 3, may comprise irrelevant data which may not be intended to be part of a measurement identifier 3 and/or may not facilitate identifying the measurement. For example, the rest of the data portion 30 hypothesized to be the measurement identifier 3 may comprise stop-words, punctuation marks and/or other words or phrases that may not be part of a measurement identifier. To alleviate this and to further facilitate the comparison between the measurement identifiers 3 and the reference data 50, the method may comprise pre-processing the measurement identifiers 3. The pre-processing is configured to bring the

detected measurement identifiers 3 in conformity with the measurement identifier references 53 thus improving the accuracy and efficiency of the comparison.

Fig. 4f illustrates the pre-processing step S32. That is, in some embodiments, wherein the method comprises detecting at least one measurement identifier 3 (i.e. step S2), the method can further comprise pre-processing the at least one detected measurement identifier 3 (i.e. step S32). The pre-processing step S32 can be performed prior to the comparison step S3. This can facilitate comparing the measurement related data 10, 20 to the reference data 50. More particularly, pre-processing at least one measurement identifier 3 can facilitate comparing the at least one measurement identifier 3 with the reference data (i.e. step S31).

The pre-processing step S32 can be advantageous as it can bring the measurement identifiers 3 comprised in the measurement related data 10, 20 in conformity with the reference data 50, more particularly, with the measurement identifier references 53. It will be understood, that the measurement identifier references 53 can also be pre-processed. For example, the pre-processing of the measurement identifier references 53 can be performed once prior to storing them (i.e. offline).

The pre-processing step S32 can be performed by the data processing system 40.

The pre-processing step S32 may comprise a data cleaning step S32A, wherein data elements 35 of the measurement identifiers 3 that do not convey information for identifying a measurement (i.e. irrelevant data elements 35', e.g. stop words 35') can be removed. The same can also be done (typically once) for the measurement identifier references 53. During the data cleaning step S32A, the data cleaning database 510 can be utilized.

The pre-processing step S32 may comprise a data element counting step S32B, wherein data elements 35 of the measurement identifiers 3 are counted. For example, the number of words of a measurement identifier 3 can be determined. Typically, the irrelevant data elements 35' are not counted. For example, the data element counting step S32B can be performed after data cleaning step S32A. Thus, each measurement identifier 3 can be associated with a corresponding data element count (e.g. word count). The same can also be done (typically once) for the measurement identifier references 53. Thus, each measurement identifier reference 53 can be associated with a corresponding data element count.

The pre-processing step S32 may comprise a data element replacement step S32C, wherein at least one data element 35 comprised by a measurement identifier 3 can be replaced with an equivalent data element 35E. For example, a word in a measurement identifier 3 can be

replaced with a respective synonym. For example, a data element 35 comprising the word "average" can be replaced with a data element 35 comprising the synonymous word "mean". Alternatively or additionally, an acronym can be replaced by the respective word(s) or phrase. During the data replacement step S32C, the data equivalent data elements database 540 can be utilized. The same can also be done (typically once) for the measurement identifier references 53. Step S32C can be advantageous as it can ensure that the measurement identifiers 3 and the identifier references 53 comprise the same data elements 35 for conveying the same information. Thus, while otherwise a comparison between a measurement identifier 3 comprising, e.g., "mean" and an identifier reference 53 comprising, e.g., "average" may erroneously yield a low similarity between the two, step S32C reduces the likelihood of such erroneous determinations (i.e. false negatives).

The pre-processing step S32 may comprise a data structure generation step S32D, wherein for each measurement identifier 3 a respective measurement identifier data structure 5 (see Fig. 6) can be generated. In such embodiments, comparing a measurement identifier 3 with the reference data 50 can comprise comparing the corresponding measurement identifier data structure 5 with the reference data 50. The measurement identifier data structure 5 can be generated using a bag-of-words model. The same can also be done (typically once) for the measurement identifier references 53. Comparing the corresponding data structures instead of comparing the measurement identifier 3 with the identifier reference 53 can increase the accuracy of results obtained from the comparison and the computational efficiency of the comparison.

The pre-processing step S32 may comprise a data-dependent value generation step S32E, wherein for each measurement identifier 3 a respective data-dependent value 7 (see Fig. 6) can be generated. The data-dependent value 7 can be generated such that similar measurement identifiers 3 can comprise similar data-dependent values 7. Moreover, the difference between the data-dependent values 7 of two measurement identifiers 3 can be proportional to the difference between the two measurement identifiers 3. The same can also be done (typically once) for the measurement identifier references 53. Similarly, the difference between the data-dependent values 7 of a measurement identifier 3 and a measurement identifier reference 53 can be proportional to the difference between the measurement identifiers 3 and the measurement identifier reference 53. Thus, each measurement identifier 3 can be associated with a respective data-dependent value 7. Similarly, each identifier reference 53 can be associated with a respective data-dependent value 7.

Fig. 4g illustrates the step of comparing a measurement identifier 3 with the reference data 50. More particularly, Fig. 4g illustrates an embodiment of the method wherein a measurement identifier 3 can be compared with the reference data 50 iteratively, wherein in each iteration of the comparison the measurement identifier 3 can be compared with a set of measurement identifier references 53. The set of measurement identifier references 53 can be selected prior to each iteration from the plurality of identifier references 53 comprised by the reference data 50. The selection is performed based on at least one predetermined criterion (e.g. number of data elements and/or data-dependent value 7 and/or code specifier 57). If during an iteration no matching is found between the measurement identifier 3 and the identifier references 53 of the set of identifier references, then a next iteration is performed. During the next iteration, another set of identifier references 53 is selected by making the selection criteria less restrictive. These steps are repeated until a matching identifier reference 53M is found for the measurement identifier 3 and/or until all the identifier references 53 of the reference data 50 are compared and/or until all the identifier references 53 of the reference data 50 comprising a code specifier 57 indicating the same code as the code 70 of the measurement related data 10, 20 (i.e. first code 70A) are compared. For the sake of brevity and to not overload Fig. 4g, the stopping condition(s) for stopping the iterative comparison (i.e. the loop) are not depicted.

The steps depicted in Fig. 4g, illustrates the comparison of one measurement identifier 3 with the reference data 50. If the measurement related data 10, 20 comprise additional measurement identifiers 3, the steps illustrated in Fig. 4g can be similarly repeated for each additional measurement identifier 3 comprised in the measurement related data 10, 20.

Moreover, if no matching identifier reference 53M can be determined from the comparison, then the measurement identifier 3 can be determined to be an uncovered measurement identifier 3. In such instances, the uncovered measurement identifiers 3 can be correspondingly labelled to indicate that no matching identifier reference 53M could be determined. Alternatively or additionally, the uncovered measurement identifiers 3 can be output, e.g., using an output user interface device, such as, a display. For example, the data processing device 40 can output the uncovered measurement identifiers 3 and a prompt requesting user input. Based on the user input, the data processing system 40 can determine the matching identifier reference 53M or can determine that the uncovered measurement identifier 3 is erroneously determined to be a measurement identifier 3 (e.g. in step S2). The user input can comprise at least one of an indication of a matching identifier reference 53M for a respective uncovered measurement identifier 3 and an indication that an uncovered measurement identifier 3 is not a measurement identifier 3. In addition, the data processing system 40, based on the user input, can add the uncovered measurement identifiers 3 to the

reference data 50, if the user input comprises an indication of a matching identifier reference 53M for a respective uncovered measurement identifier 3. Moreover, the data processing system 40, based on the user input, can add the uncovered measurement identifiers 3 to the reference data 50 associated with the indicated matching identifier reference 53M.

In a step S310, comparing a measurement identifier 3 with the reference data 50 can comprise selecting a set of measurement identifier references 53 from the reference data 50. The selection in step S310 can be performed based on at least one selection criteria. The at least one selection criteria can be configured such that the measurement identifier references 53 with the highest likelihood of matching to the measurement identifier 3 can be selected. For example, the selection criteria in step S310 can comprise selecting only the measurement identifier references 53 that comprise the same data element count as the measurement identifier 3. Alternatively, the selection criteria in step S310 can comprise selecting only the measurement identifier references 53 that comprise a data element count within a first data element count range centered at the data element count of the measurement identifier (i.e. the absolute value of the difference between the data element count of the measurement identifier references 53 and the data element count of the measurement identifier 3 is not larger than a first data element count threshold). Alternatively or additionally, the selection criteria in step S310 can comprise selecting only the measurement identifier references 53 that comprise a data-dependent value within a first data-dependent value range centered at the data-dependent value of the measurement identifier 3 (i.e. the absolute value of the difference between the data-dependent value of the measurement identifier references 53 and the data-dependent value of the measurement identifier 3 is not larger than a first data-dependent value threshold). Alternatively or additionally, the selection criteria in step S310 can comprise selecting only the measurement identifier references 53 that correspond to the same code 70 as the measurement related data 10, 20.

In a step S311 the measurement identifier 3 can be compared with each measurement identifier reference 53 comprised in the selected set.

In a step S312, during each comparison a similarity score can be calculated. Thus, for each measurement identifier reference 53 in the selected set a respective similarity score can be calculated based on the comparison with the detected measurement identifier 3. The similarity score can be a number calculated to proportionally indicate similarity.

At a first iteration, in step S311 and S312 the selected set corresponds to the set selected during step S310.

In a step S313, each similarity score can be compared with a matching threshold. Thus, in step S313 it can be determined whether at least one of the similarity scores is larger than or equal to the matching threshold. It will be understood that this is the case when the similarity score proportionally indicates similarity. If there is at least one similarity score larger than or equal to the matching threshold, then the comparison step S3 can terminate with step S315 wherein at least one matching measurement identifier reference 53M can be determined. More particularly, the measurement identifier 3 can be matched with the measurement identifier reference(s) 53, with which (a) similarity score(s) larger than or equal to the matching threshold is/are calculated in step S312.

Alternatively, if in step S313 it can be determined that all the similarity scores calculated in step S312 are smaller than the matching threshold, then the comparison step may re-iterate through step S314 to step S311. In step S314, another set of measurement identifier references can be selected from the reference data 50. For example, in step S314 the selection criteria used in step S310 can be made less restrictive. For example, the data count range can be extended (i.e. the data element count threshold increased) and/or the data-dependent value range can be extended (i.e. the data-dependent value threshold can be increased) and/or identifier references 53 corresponding to codes other than the code of the measurement related data 10, 20 can be considered. This can allow for other measurement identifier references 53 to be selected (and thus compared with the measurement identifier 3). It will be understood, that in step S314 previously selected and compared identifier references 53 are not selected again in the set of measurement identifier references 53 – although they may fulfil the less-restrictive selection criteria. That is, the measurement identifier references 53 are compared only once with the measurement identifier 3.

Next, in step S311 (second iteration), the newly selected measurement identifier references 53 of the set selected during the previous iteration can be compared with the measurement identifier. Next, steps S312 and S313 can be performed as discussed.

Thus, the comparison step can be performed iteratively. This can be advantageous as it can decrease the number of comparisons required to find a matching measurement identifier reference 53M. That is, performing the comparison step iteratively as discussed above can allow initially comparing the measurement identifier 3 with measurement identifier references 53 that comprise a high likelihood of matching to the measurement identifier 3. Thus, the measurement identifier 3 can be matched faster and with a measurement identifier reference 53. Additionally, less computational resources may be required.

Fig. 5 graphically illustrates the steps depicted in Fig. 4g. More particularly, Fig. 5 illustrates the embodiment wherein the selection criteria used in steps S310 and S314 are based on the data element count value and data-dependent value of the measurement identifier 3 and the measurement identifier references 53. It will be understood that this is an example and that the selection criteria may consider further parameters, such as, the code of the measurement related data 10, 20 and the code specifier 57 corresponding to each identifier reference 53.

The horizontal axis depicts the data element count value and the vertical axis depicts the data-dependent value. The filled circle indicates the measurement identifier 3 that is being compared plotted according to its respective data element count (see step S32B, Fig. 4f) and its respective data-dependent value (see step S32E, Fig. 4f). The other markers (crosses, empty circles, triangles and pentagons) correspond to the measurement identifier references 53 comprised by the reference data 50 plotted according to their respective data element counts and data dependent values.

The vertical filled line indicates the boundary for selecting the first set of measurement identifier references 53 (step S310). In this example, the first set is selected to include measurement identifier references 53 that comprise same the data element count value with the measurement identifier 3 and a respective data-dependent value within a range (indicated by the length of the line) centered at the data-dependent value of the measurement identifier 3. During the first iteration, the measurement identifier 3 is compared with all measurement identifier references 53 that fulfil these criteria (i.e. the ones plotted with crosses).

The dotted square (i.e. square with dotted lines) indicates the boundary for selecting the second set of measurement identifiers (i.e. first execution of step S314). As indicated, the selection criteria are made less restrictive, i.e., the data element count range and the data-dependent value range are increased. During the second iteration, the measurement identifier 3 can be compared with the newly selected measurement identifier references 53 (plotted as empty circles). It will be noted that the measurement identifier references 53 selected during the first iteration (plotted as crosses) are not compared again during the second iteration (although they are within the boundaries). Moreover, it will be understood that the second iteration is performed only if none of the measurement identifier references 53 selected during the first iteration (plotted as crosses) matches the measurement identifier 3.

The dashed square (i.e. square with dashed lines) indicates the boundary for selecting the third set of measurement identifiers during the third iteration (i.e. second execution of step S314). As indicated, the selection criteria are made even less restrictive, i.e., the data element count range and the data-dependent value range are increased. During the third iteration, the

measurement identifier 3 can be compared with the newly selected measurement identifier references 53 (plotted as triangles). It will be noted that the measurement identifier references 53 selected during the first and second iteration are not compared again during the third iteration (although they are within the boundaries). Moreover, it will be understood that the third iteration is performed only if none of the measurement identifier references 53 selected during the second iteration matches the measurement identifier 3.

During the third iteration, it can be found that one of the measurement identifier references 53 (plotted with a filled triangle) matches the measurement identifier 3. Thus, the comparison method may terminate (step S315).

Fig. 6 graphically illustrates the method steps discussed above. As depicted, the obtained measured related data can be processed such that at least one measurement identifier 3 can be detected (Figs. 4c to 4e). For example, at least one data portion 30 and at least one data element 35 can be detected. Moreover, for each detected data portion 30 it can be determined whether it comprises a measurement identifier 3. For example, each data element 35 can be classified as corresponding to a measurement indicator 2 or to a measurement identifier 3.

Further, each detected measurement identifier 3 can be pre-processed (Fig. 4f). This can comprise removing irrelevant measurement identifiers 35', replacing at least one data element 35 with an equivalent data element 35E, generating a measurement identifier data structure 5, generating a data element count and/or generating a data-dependent value 7.

After pre-processing the measurement identifier 3 (and/or the associating data generated during pre-processing) can be compared with reference data 50. Thus, a matching measurement identifier reference 53M can be determined.

Moreover, based on the matching identifier reference 53M the method can comprise determining a replacing identifier reference 53R corresponding to the measurement identifier 3. If the matching identifier reference 53M corresponds to the second code 70B (i.e. to the code 70 according to which the measurement related data 10, 20 are to be configured), then the replacing identifier reference 53R is the matching identifier reference 53R. Otherwise, an identifier reference 53 comprised by the reference data 50, configured according to the second code 70B and associated with the matching identifier 53M can be determined to be the replacing identifier reference 53R.

The data processing system 40 can be configured to replace each measurement identifier 3 with the corresponding replacing identifier reference 53R. Thus, the data processing system 40 can configure the measurement related data 10, 20 according to the second code 70B.

In addition, the data processing system 40 can replace at least one identifier indicator 2 corresponding to a measurement identifier 3, with a reference characteristic 55 corresponding to the replacing identifier reference 53R.

Fig. 7 indicates the measurement related data 10, 20 of Fig. 2, wherein the measurement identifiers 3 are replaced with the respectively replacing identifier reference 53R. That is, Fig. 7 illustrates the measurement related data 10, 20 configured according to the second code 70B.

Fig. 8a illustrate a portion of the reference data 50 corresponding to an exemplary measurement. More particularly, Fig. 8a depicts a plurality of measurement identifier references 53 corresponding to a measurement and configured according to an intermediate code 70I, which can be a standard code 70I. In the depicted example, the measurement is Cholesterol HDL measurement and the intermediate code 70I is the LOINC database. That is, Fig. 8a depicts code names 53 and long names 53 of the LOINC code 70I corresponding to the measurement of cholesterol HDL. As illustrated, the reference data 50 can comprise a plurality of measurement identifier references 53 corresponding to the LOINC database 70I (each comprising a number and a long common name). In addition, each measurement identifier reference 53 is further associated with a plurality of reference characteristics 55. In the depicted example, the reference characteristics 55 further characterise the component to be measured, the property (i.e. unit), the system (i.e. component on which the measurement is to be performed), the scale type (e.g. quantitative) and the method type (e.g. electrophoresis).

As discussed, in addition to standard (or intermediate) measurement identifier references 53 the reference data 50 can comprise further identifier references 53 which can be non-standard or non-universal. Fig. 8b, illustrates a plurality of non-standard biomarkers 53 (e.g., synonyms, short names, acronyms) corresponding to the Cholesterol HDL measurement. Each of the further identifier reference 53 can be associated with a corresponding code specifier 57 indicating a code that the corresponding identifier reference 53 belongs to. For example, the code specifier 57 can indicate which node (e.g. which laboratory, hospital or physician) uses the respective identifier reference 53 to refer to the measurement. In addition, the reference data 50 can comprise other possible reference characteristics 55 (e.g. other possible units 55) that can be used for the measurement (e.g. for the Cholesterol HDL measurement).

Fig. 8c illustrates an example of measurement related data 10, 20 that can be generated by a medical laboratory. The measurement related data 10, 20 relate to a plurality of measurements. Each measurement (depicted in a respective row of the table) is identified by a measurement identifier 3, i.e., by a biomarker 3, which is used internally by the respective laboratory. As illustrated, the measurement identifiers 3 can comprise short names, acronyms and/or particular phrases to refer to measurements. In addition, the measurement related data 10, 20 can comprise fields (e.g. cells of the table) which can be filled with measurement indicators 2 to further characterize the respective measurements they correspond to. In the depicted example, only the measurement identifiers 3 are provided for each measurement.

It will be understood, that the measurement related data 10, 20 are depicted in the tabular form in Fig. 8c to facilitate the visualization and illustration of the measurement related data 10, 20. In general, the measurement related data 10, 20 may be electronically stored in a memory device using an adequate data structure (e.g. as illustrated in Fig. 2).

Fig. 8d depicts a plurality of matching measurement identifier references 53M that can be determined for the "HDL chol." Measurement identifier. Associated with each matching measurement identifier references 53M the respective calculated similarity score 9 can be output. As illustrated in the depicted example, the measurement identifier can be matched with a plurality of measurement identifier references 53. In some embodiments, the matching measurement identifier references 53M with the highest similarity score 9 can be considered. Alternatively or additionally, if the measurement in the measurement identifier reference is further characterized (i.e. by providing at least one measurement indicator 2) it can be used to filter the plurality of matching measurement identifier references 53M. For example, if in the measurement related data 10, 20 of Fig. 8c, the Method_Typ characteristic 55 would be specified as "Electrophoresis", then the identifier reference "49130-8" can be considered as the matching identifier reference 53M. Alternatively or additionally, the plurality of matching measurement identifier references 53M can be output to a user and based on user input one of the plurality of matching measurement identifier references 53M can be considered (while the others disregarded).

Fig. 9a depicts an example of measurement related data 10, 20 configured according to a first code, i.e., measurement related data 10, 20 that can be obtained by the data processing system 40 (see, e.g., Fig. 1). More particularly, in Fig. 9a an image of a laboratory report 10, 20 is depicted. Fig. 9a illustrates a typical input that can be provided to the present invention, e.g., to the data processing system 40. That is, Fig. 9a depicts typical measurement related data 10, 20 that can be processed by the present invention.

As indicated, the measurement related data (e.g. the laboratory report) 10, 20 can comprise a plurality of data elements 35. For example, each data element 35 can be a word in the laboratory report 10, 20. It will be noted that not to overload the Figure, some referral numbers are omitted and not shown in Fig. 9a. In addition, the measurement related data (e.g. the laboratory report) 10, 20 can comprise a plurality of data portions 30. For example, each data portion 30 can be a line in the laboratory report 10, 20. Some of the data portions 30 can correspond to measurements, i.e. can comprise at least one measurement identifier 3, e.g., a biomarker. For example, lines 2 to 16 are biomarker lines, i.e., data portions 30 comprising at least one measurement identifier 3. However, some data portions 30 may not correspond to measurements. For example, line 1 is a non-biomarker line, i.e., a data portion 30 that does not comprise a measurement identifier 3.

Moreover, the measurement related data (e.g. the laboratory report) 10, 20 can comprise measurement identifiers 3. For example, column 1 lists a plurality of biomarkers 3, each provided in a respective row.

In addition, the measurement related data (e.g. the laboratory report) 10, 20 can comprise measurement indicators 2. For example, in columns 2 to 5, the laboratory report comprises for each biomarker 3 a plurality of respective measurement indicators 2. More particular, column 2 comprises values of the measurements, column 3 comprises units, column 4 comprises ranges and column 5 comprises methods.

As illustrated, the measurement data 10, 20 can comprise an image. In such embodiments, the present invention can be configured to process the image and obtain text data based on the processing of the images.

Moreover, the present invention can be configured to determine the structure of the measurement related data 10, 20. This can be particularly advantageous when or if the measurement related data comprise at least one image and/or are provided as image data. For example, the present invention can be configured to detect, on the measurement related data 10, 20, data portions 30, data elements 35, measurement identifiers 3 and/or measurement indicators 2. Furthermore, when detecting measurement indicators 2, the present invention can be configured to detect units, numerical values, text values and/or ranges.

Fig. 9b illustrates an exemplary output after processing the measurement related data 10, 20 of Fig. 9a. Fig. 9b can be an example of the measurement related data 10, 20 configured

according to the second code. For the sake of brevity, only the output corresponding to the second line of the laboratory report of Fig. 9a is depicted.

As illustrated, the measurement related data 10, 20 can be output (e.g. by the data processing system) in a structured form, e.g., in a tabular form. Different parts of the measurement related data 10, 20 can be labeled. For example, the data elements 35 corresponding to the measurement identifier 3 can be labeled. In the provided example, the biomarker "TOTAL LEUCOCYTES COUNT" is labeled as "Biomarker in Report". Similarly, the value 2 of "5.08" and the unit in the report 2 " $\times 10^3/\mu\text{L}$ " are correspondingly labeled. Moreover, the measurement identifiers 3 and the corresponding measurement indicators 2 can be associated together, e.g., by arranging them in the same row of a table. It will be noted that this association may not be readily extractable from the original measurement related data 10, 20 – particularly when the measurement related data 10, 20 can be provided as image data, as illustrated in Fig. 9a.

In addition to the data extracted from the original measurement related data 10, 20, the present invention can output the identifier references 53 that can be matched with the measurement identifiers 3 comprised in the measurement related data 10, 20. This can be advantageous as it can allow outputting the measurement related data 10, 20 configured according to the second code. As illustrated the identifier references "WBC" and "Leucocyte Count" can be matched with the measurement identifier "TOTAL LEUCOCYTES COUNT" from the original measurement related data 10, 20. The identifier references "WBC" is a biomarker abbreviation and the identifier references "Leucocyte Count" is a biomarker name. The identifier references 53 "WBC" and "Leucocyte Count" matched with the measurement identifier 5 "TOTAL LEUCOCYTES COUNT", can for example correspond to the second code (i.e. to the target code).

Furthermore, the reference characteristics 55 (e.g. biomarker Unit) corresponding to each identifier reference 53 respectively matched with the measurement identifiers 3 from the measurement related data 10, 20 can be output. This can further facilitate outputting the measurement related data 10, 20 configured according to the second code. For example, the unit 55 of the identifier reference "WBC" and "Leucocyte Count" can be output in the same row. In addition, the range and the boundaries of the range (see Min, Max) can be output. Similarly, the unit 2 in the report " $\times 10^3/\mu\text{L}$ " can be mapped to the standard unit 55 " $10^3/\mu\text{L}$ ".

Thus, the data from measurement related data 10, 20 (which can be in an image format) can be matched with structured data (i.e. with identifier references). The identifier references 53 can be configured according to known, standard and/or universal standards. Measurement

related data 10, 20, such as, the one illustrated in Fig. 9a, may allow a very limited use. For example, the laboratory report of Fig. 9a, which is an image, may only be read by a human, e.g., a physician and, more particularly, by a physician which is familiar with the biomarker names used in the laboratory report. For example, a physician not familiar with the abbreviation "RBC" may not understand the last line of the laboratory report of Fig. 9a.

The present invention can process the obtained measurement related data 10, 20 and can output structured data corresponding to the obtained measurement related data 10, 20. After processing, the measurement related data 10, 20 can be read not only by a human but can also be more easily decoded automatically by processing units. For example, a laboratory information system (LIS) can easily process the output of the present invention. This can be particularly advantageous, as typically, the computational resources in a laboratory, clinic or a physician's office may be limited (e.g. a laptop, tabled and/or a PC). As such, an automatic processing of the measurement related data 10, 20 in their original form may not be feasible. However, by structuring the measurement related data 10, 20 the present invention alleviates the need of complex computations to handle the measurement related data 10, 20.

The latter is further facilitated by configuring the measurement related data 10, 20 according to known codes and/or standards. For example, the present invention can map the biomarker names of a laboratory report with standard biomarker names (e.g. with LOINC biomarker names). In another example, the measurement related data 10, 20 can be configured, after processing, according to a code that is known by the receiving node, e.g., a laboratory information system to which the measurement related data will be provided. For example, the measurement related data can be structured in a form known or expected by the receiving node. For example, the measurement related data can be structured in a tabular form as illustrated in Fig. 9b. Furthermore, the measurement identifiers, units, ranges, etc., can be output such that they match with the measurement identifiers, units, ranges, etc., used by the receiving node. Thus, the receiving node may readily obtain information from the measurement related data.

Fig. 10 illustrates units that can be comprised by the data processing system 40. The units 401 to 415 can implement the steps of the method discussed above, e.g., see Figs. 4a to 4g.

In some embodiments, at least one of the units 401 to 415 can be implemented in software (i.e. can be a computer program) and can be executed by the data processing system 40 (see Fig. 1). The at least one processing unit 401 to 415 implemented in software may also be referred to as a processing program. For example, all the units 401 to 415 can be implemented in software and can be executed by a general central processing unit.

In some embodiments, at least one of the units 401 to 415 can be implemented in hardware. For example, at least one of the units 401 to 415 can comprise at least one of integrated circuit, central processing unit (CPU), graphical processing unit (processing unit) GPU, digital signal processor (DSP), accelerated processing unit (APU), application specific integrated circuit (ASIC), application specific instruction set processor (ASIP), field programmable gate array (FPGA), artificial intelligence (AI) accelerator and tensor core, each of which can be in the singular or plural.

In some embodiments, at least one of the units 401 to 415 can be implemented as a mixture between hardware and software components. For example, at least one of the units 401 to 415 can comprise a software component loaded and a hardware component configured to execute the software component.

The units 401 to 415 can be integrated into and/or executed by the data processing system 40 – see Fig. 1. That is, the data processing system can comprise any of, preferably all, the units 401 to 415.

In addition, Fig. 10 depicts a database comprising the reference data 50.

An input unit 401 can be configured to receive measurement related data 10, 20. For example, the input unit 401 can receive the measurement related data 10, 20 from a sending node in a communication network. The obtained measurement related data 10, 20 can be configured according to a first code. For example, the measurement related data 10, 20 can be a laboratory report comprising image data as illustrated in Fig. 9a. The input unit 401 can for example carry out step S1 of the method illustrated in Fig. 4a.

The obtained measurement related data 10, 20 can be provided to an entities recognition unit 403. The entities recognition unit 403 can be configured to detect data elements 35 and/or data portions 30 that can be comprised by the obtained measurement related data 10, 20. This can be particularly advantageous if the measurement related data 10, 20 can comprise image data. Moreover, the entities recognition unit 403 can be configured to detect measurement identifiers 3 and/or measurement indicators 2. For example, the entities recognition unit 403 can be configured to carry out step S2 of the method illustrated in Fig. 4b. More particularly, the entities recognition unit 403 can be configured to carry out any of the methods illustrated in Figs. 4c to 4e.

For example, the entities recognition unit 403 can loop through lines 30 of a laboratory report 10, 20 and can classify each line as non-biomarker line, range data line and biomarker line.

Non-biomarker lines can be report lines that do not comprise biomarker data, e.g., lab info, address, patient profile data. The range data lines can be report lines that contain ranges data. The entities recognition unit 403 can detect normal range borders and can associate this info with the relevant biomarker. The biomarker lines can be report lines that contain some biomarker data, e.g., biomarker name, value, text value, unit, range. To recognize entities in biomarker lines or ranges lines, the entities recognition unit 403 can utilize a set of predefined regular expressions.

The detected measurement identifiers 3 can be provided directly to a matching unit 409 after being detected by the entities recognition unit 403. However, in preferred embodiments, the detected measurement identifiers 3 can be provided directly to an online pre-processing unit 405. The online pre-processing unit 405 can be configured to pre-process the detected measurement identifiers 3. As discussed, the pre-processing step can increase the homogeneity between the measurement identifiers 3 and the reference data 50, thus facilitating their comparison. The online pre-processing unit 405 can be configured to carry out any of the steps depicted in Fig. 4f.

For example, the online pre-processing unit 405 can be configured to carry out the following pre-processing steps:

A data cleaning step which can comprise a stop-words removal step.

A synonym replacement step, wherein a synonyms dictionary can be utilized. Replacing synonyms may be advantageous, as it may facilitate recognition of identical or similar names.

For example, when removing stop words and replacing synonyms from "The High-Density Lipid Cholesterol" and "Cholesterol (High Density Lipoprotein)", a same term can result (only with a different order of words).

Converting plural form to singular which may be advantageous, as it may allow for more precise comparison of words, particularly if some words are indicated in plural, and others in singular. For example, when converting plural forms to singular and replacing synonyms from "Mean Cell Volume" and "Cells average size", a same term can result (only with a different order of words).

A data element counting step, wherein, e.g., for each measurement identifier the data elements can be counted (preferably after the data cleaning step) and the relevant counter can be associate to the measurement identifier.

A data dependent value generation step, which may be performed by means of a hash function, which may be cryptographic. Generating a data-dependent value (e.g., a hash-value) for each of the names may be advantageous as it can allow to easily detect terms that may be identical, since their hash-value may be identical. Further, comparing hash values may be less computing-time intensive than (literally) comparing the entire names. That is, it can be less computationally complex to compare the respective data-dependent values than directly comparing the measurement identifiers with the identifier references.

Furthermore, some hash functions, such as a sum of parts of a name, may generate a same result for two elements that are identical apart from an exchanged order. This can be a resource-efficient way to estimate names or elements that can be identical apart from an interchange of parts, i.e. a changed order of the parts, compared to a comparison algorithm that exactly determines whether two identifier or elements are identical apart from the interchange.

It will be understood, that the same pre-processing operations can also be performed on the identifier references 53 of the reference data 50. More particularly, an offline pre-processing unit 413 can pre-process the measurement identifiers. The offline pre-processing unit 413 can be identical or the same to the online pre-processing unit 405.

In some embodiments, the offline pre-processing unit 413 may not be part of the data processing system 40.

A matching unit 409 can be configured to compare the measurement related data 10, 20 with the reference data 50. More particularly, the matching unit can be configured to match each of the measurement identifiers 3 of the measurement related data 10, 20 with at least one identifier reference 53. The matching identifier 409 can thus obtain the measurement related data 10, 20 and particularly the detected and preferably pre-processed measurement identifiers 3. In some embodiments, only the detected and preferably pre-processed measurement identifiers 3 can be provided to the matching unit 409. In some embodiments, the detected and preferably pre-processed measurement identifiers 3 and the corresponding measurement indicators 2 can be provided to the matching unit 409. In some embodiments, the entire data portion 30 comprising the measurement identifier 3 can be provided to the matching unit 409. The matching unit 409 can be configured to determine and further optionally to output the matching identifier reference 53M and/or replacing identifier reference 53R (see Fig. 6). The matching unit 409 can be configured to carry out step S3 of the method of Fig. 4a, preferably step S31 of the method of Fig. 4b, further preferably steps S310 to S315 of the method of Fig. 4g.

The main objective of the matching unit 409 can be to match each measurement identifier of the measurement related data 10, 20 with the most appropriate (similar) identifier reference 53 from the reference data 50.

For example, for each report biomarker line (i.e. for each data portion 30 comprising a measurement identifier 3) the following operations can be performed:

As discussed, the pre-processing units 405, 4013 can pre-process the measurement identifiers 3 and the identifier references 53. Each measurement identifier 3 and each identifier reference 53 can be associated with a respective data-dependent (e.g. hash) value and a respective data element count.

The matching unit 409 can calculate progressively the similarity between each measurement identifier and the identifier references 53 comprised by the reference data 50. That is, instead of "blindly" comparing with all the identifier references 53 heuristic rules can be used to minimize the necessary execution time of the matching algorithm. More particularly, the matching unit 409 can be configured to initially compare the measurement identifier 3 with identifier references 53 that can be similar to the measurement identifier 3. For example, the matching unit can start by comparing the measurement identifier 3 with identifier references 53 that have approximately the same count of data elements and a data-dependent value within some interval which center can be the data-dependent value of the measurement identifier 3. Further, the matching unit 409 can incrementing the interval until matching succeeds. If no matching is found, the matching unit 409 can considers remaining identifier references 53. The identifier reference with the maximum similarity can be considered as the resulting matching identifier reference 53M (an internal symbol may be associated).

The matching unit 409 can utilize a similarity metric calculating unit 407 to calculate a similarity metric between a measurement identifier 3 and an identifier reference 53.

The similarity metric calculating unit 407 can be configured to receive two inputs and calculate a similarity metric between the two inputs. For example, the similarity metric calculating unit 407 can be configured to receive a multiset data structure corresponding to the measurement identifier 3 and a multiset data structure corresponding to the identifier reference 53. To calculate the similarity between the two inputs, a Metaphone algorithm can be used to calculate the maximum similarity ratio for each word of the first input with a relevant word of the second input. Such a ratio may for example be 0.91 for a comparison of "this" and "these". For example, the similarity metric calculating unit 407 may calculate the similarity metric using

$$X = \frac{2 \cdot \sum_i ratio_i}{n_1 + n_2}$$

wherein $ratio_i$ is the above-mentioned ratio of the comparison for each word, the index i indicates an ordinal number of said word, and n_1 and n_2 indicate amounts of words of the inputs to which the comparison criterion is applied and X indicates the similarity metric.

For example, the inputs "How close is this to that" and "these two are not that close" can be provided to the pre-processing units 405, 413 for generating the following multisets:

$s1 = ['how', 'close', 'is', 'this', 'to', 'that']$

$s2 = ['these', 'two', 'are', 'not', 'that', 'close']$.

The two multisets can be provided to the similarity metric calculating unit 407 which can calculate the similarity metric X as the sum of (0.27 (how/not), 1(close/close), 0(is/-), 0.9(this/these), 0.3 (to/two), 1(that/that) divided by 6. In this exemplary case, the comparison criterion may thus yield 0.57.

In addition, a learning unit 411 can be provided. The learning unit 411 can be configured to extend the reference data 50 with further identifier references 53 and/or measurement characteristics 55. For example, the learning unit 411 can receive a user input which can comprise identifier references 53, units corresponding to an identifier reference, new units being synonyms or equivalent to existing units and/or new identifier references 53 being synonyms or equivalent to existing identifier references 53. Based on the user input, the learning unit 411 can extend the reference data 50.

In some instances, the matching unit 409 can be inconclusive. That is, the matching identifier 409 may not find a matching identifier reference 53 for a measurement identifier 3. In some embodiments, the unmatched measurement identifier 3 can be provided to the learning unit 411 – which can for example further display it to a user. Based on user input, a matching identifier reference 53M can be determined for the unmatched measurement identifier 3. The learning unit 411 can add the unmatched measurement identifier 3 to the reference data 50 and can associate the unmatched measurement identifier 3 with the matching identifier reference 53M. Thus, if the unmatched measurement identifier 3 is detected in a future measurement related data, it can be matched automatically by the matching unit 409 to the respective matching identifier reference 53M.

Whenever a relative term, such as “about”, “substantially” or “approximately” is used in this specification, such a term should also be construed to also include the exact term. That is, e.g., “substantially straight” should be construed to also include “(exactly) straight”.

It should also be understood that whenever reference is made to an element this does not exclude a plurality of said elements. For example, if something is said to comprise an element it may comprise a single element but also a plurality of elements.

Whenever steps were recited in the above or also in the appended claims, it should be noted that the order in which the steps are recited in this text may be accidental. That is, unless otherwise specified or unless clear to the skilled person, the order in which steps are recited may be accidental. That is, when the present document states, e.g., that a method comprises steps (A) and (B), this does not necessarily mean that step (A) precedes step (B), but it is also possible that step (A) is performed (at least partly) simultaneously with step (B) or that step (B) precedes step (A). Furthermore, when a step (X) is said to precede another step (Z), this does not imply that there is no step between steps (X) and (Z). That is, step (X) preceding step (Z) encompasses the situation that step (X) is performed directly before step (Z), but also the situation that (X) is performed before one or more steps (Y1), ..., followed by step (Z). Corresponding considerations apply when terms like “after” or “before” are used.

While in the above, a preferred embodiment has been described with reference to the accompanying drawings, the skilled person will understand that this embodiment was provided for illustrative purpose only and should by no means be construed to limit the scope of the present invention, which is defined by the claims.

Furthermore, reference numbers and letters appearing between parentheses in the claims, identifying features described in the embodiments and illustrated in the accompanying drawings, are provided as an aid to the reader as an exemplification of the matter claimed. The inclusion of such reference numbers and letters is not to be interpreted as placing any limitations on the scope of the claims.

Claims

1. A data processing system (40) comprising
an input unit (401) configured to obtain measurement related data (10, 20) configured according to a first code (70A);
a matching unit (409) configured to
compare the obtained measurement related data (10, 20) with reference data (50) and
to
configure the measurement related data (10, 20) according to a second code (70B)
based on the comparison.
2. The data processing system (40) according to the preceding claim, further comprising
an entities recognition unit (403) configured to detect in the obtained measurement related data (10, 20) at least one of
a data element (35), a data portion (30), a measurement identifier (3) and a measurement indicator (2).
3. The data processing system (40) according to the preceding claim, wherein the data processing system (40) comprises an online pre-processing unit (405) configured to pre-process the at least one detected measurement identifier (3).
4. The data processing system (40) according to any of the preceding claims, wherein the data processing system (40) comprises a similarity metric calculating unit (407) configured to receive two inputs and to calculate a similarity metric between the first input and the second input and
wherein the matching unit (409) is configured to utilize the similarity metric calculating unit (407).
5. A method comprising:
a data processing system (40) obtaining measurement related data (10, 20) configured according to a first code (70A);
the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50);
based on the comparison, the data processing system (40) configuring the measurement related data (10, 20) according to a second code (70B).

6. The method according to the preceding claim,
wherein the obtained measurement related data (10, 20) comprises at least one measurement identifier (3) configured to identify a measurement and
wherein the method comprises detecting the measurement identifier (3) in the obtained measurement related data (10, 20).
7. The method according to the preceding claim,
wherein the obtained measurement related data (10, 20) comprises a plurality of data portions (30),
wherein the method comprises detecting at least one data portion (30) in the obtained measurement related data (10, 20), and
wherein detecting the measurement identifier (3) comprises detecting the data portion (30) of the obtained measurement related data (10, 20) that comprises the measurement identifier (3).
8. The method according to any of the 2 preceding claims, wherein the method comprises detecting at least one measurement indicator (2) in the obtained measurement related data (10, 20),
wherein the at least one measurement indicator (2) comprises at least one of numerical values, ordinal values, nominal values, qualitative data, quantitative data, unit of measurement, range and range specifier and
wherein detecting the measurement identifier (3) comprises detecting at least one measurement identifier (3) based on the detection of at least one measurement indicator (2).
9. The method according to the preceding claim, wherein the method comprises determining the location of the measurement indicator (2) in the obtained measurement related data (10, 20) and
wherein detecting the measurement identifier (3) comprises detecting at least one measurement identifier (3) based on the location of the at least one measurement indicator (2) in the obtained measurement related data (10, 20).
10. The method according to any of the 3 preceding claims, wherein the method comprises determining whether the at least one data portion (30) comprises at least one measurement indicator (2) and
wherein detecting at least one data portion (30) of the obtained measurement related data (10, 20) that comprises the measurement identifier (3) comprises

determining that at least one data portion (30) comprises the measurement identifier (3) if the data portion (30) comprises the measurement indicator (2).

11. The method according to any of the 5 preceding claims, wherein each of the measurement identifiers (3) comprises at least one data element (35),

wherein the method comprises pre-processing the detected measurement identifier (3) to facilitate the step of the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50),

wherein pre-processing the detected measurement identifier (3) comprises

counting the number of data elements (35) comprised by the detected measurement identifier (3) and

generating for each detected measurement identifier (3) a corresponding data dependent value (7), wherein the data dependent value (7) corresponding to a measurement identifier (3) depends on the data comprised by the measurement identifier (3).

12. The method according to any of the 6 preceding claims,

wherein the reference data (50) comprises a plurality of measurement identifier references (53),

wherein the data processing system (40) comparing the obtained measurement related data (10, 20) with reference data (50) comprises comparing each detected measurement identifier (3) with the reference data (50) and

wherein the method comprises determining for each detected measurement identifier (3), respectively, a matching measurement identifier reference (53M) based on the comparison between the measurement identifier (3) and the reference data (50).

13. The method according to the preceding claim, wherein each comparison between each detected measurement identifier (3) with the reference data (50) is performed iteratively,

wherein during each iteration of each comparison between each detected measurement identifier (3) with the reference data (50) the method comprises

determining a set of measurement identifier references (53), wherein the set of measurement identifier references (53) is a sub-set of the plurality of measurement identifier references (53) of the reference data (50) and

comparing the respective detected measurement identifier (3) with each of the measurement identifier references (53) comprised by the set of measurement identifier references (53) determined during that iteration.

14. The method according to the preceding claim and with the features of claim 11, wherein determining a set of measurement identifier references (53) during each iteration comprises selecting the set of measurement identifier references (53) out of the plurality of measurement identifier references (53) comprised by the reference data (50), wherein selecting the set of measurement identifier references (53) out of the plurality of measurement identifier references (53) comprised by the reference data (50) comprises selecting each measurement identifier reference (53) if the number of data elements of the measurement identifier reference (53) is within a data-element-count range corresponding to that iteration and if the data dependent value (7) of the measurement identifier reference (53) is within a data-dependent-value range corresponding to that iteration.

15. The method according to any of the 2 preceding claims, wherein during each iteration of each comparison between each detected measurement identifier (3) with the reference data (50) the method comprises calculating a respective similarity metric between the respective detected measurement identifier (3) and each of the measurement identifier references (53) comprised by the set of measurement identifier references (53) determined during that iteration, comparing each calculated similarity metric with a matching threshold, determining at least one matching measurement identifier reference (53M) depending on the comparison of each of the calculated similarity metrics with the matching threshold.

16. The method according to any the preceding claims 5 to 15, wherein configuring the measurement related data (10, 20) according to the second code (70B) comprises determining for each measurement identifier (3) a replacing measurement identifier reference (53R), replacing each measurement identifier (3) with the respective replacing measurement identifier reference (53R), wherein the replacing measurement identifier reference (53R) is a measurement identifier reference (53) configured according to the second code (70B)

wherein determining for each measurement identifier (3) a replacing measurement identifier reference (53R) depends on the respective matching measurement identifier reference (53M) determined for the respective measurement identifier (3).

17. The method according to any of the preceding claims 5 to 16, wherein the method comprises

- a measurement requesting node
 - generating measurement instruction data (10) configured according to the first code (70A) and
 - sending the measurement instruction data (10) to the data processing system (40);
- the data processing system (40) sending the measurement instruction data (10) configured according to the second code (70B) to a measurement performing node;
- the measurement performing node
 - performing the requested measurement(s),
 - generating measurement result data (20) configured according to the second code (70B), and
 - sending the measurement result data (20) to the data processing system (40);
- the data processing system configuring the measurement result data (20) according to the first code (70A) and sending the measurement result data (20) to the measurement requesting node.

18. The method according to any of the preceding claims 5 to 17, wherein the method is a computer-implemented method.

19. A communication system comprising a data processing system (40) wherein the data processing system (40) is configured to

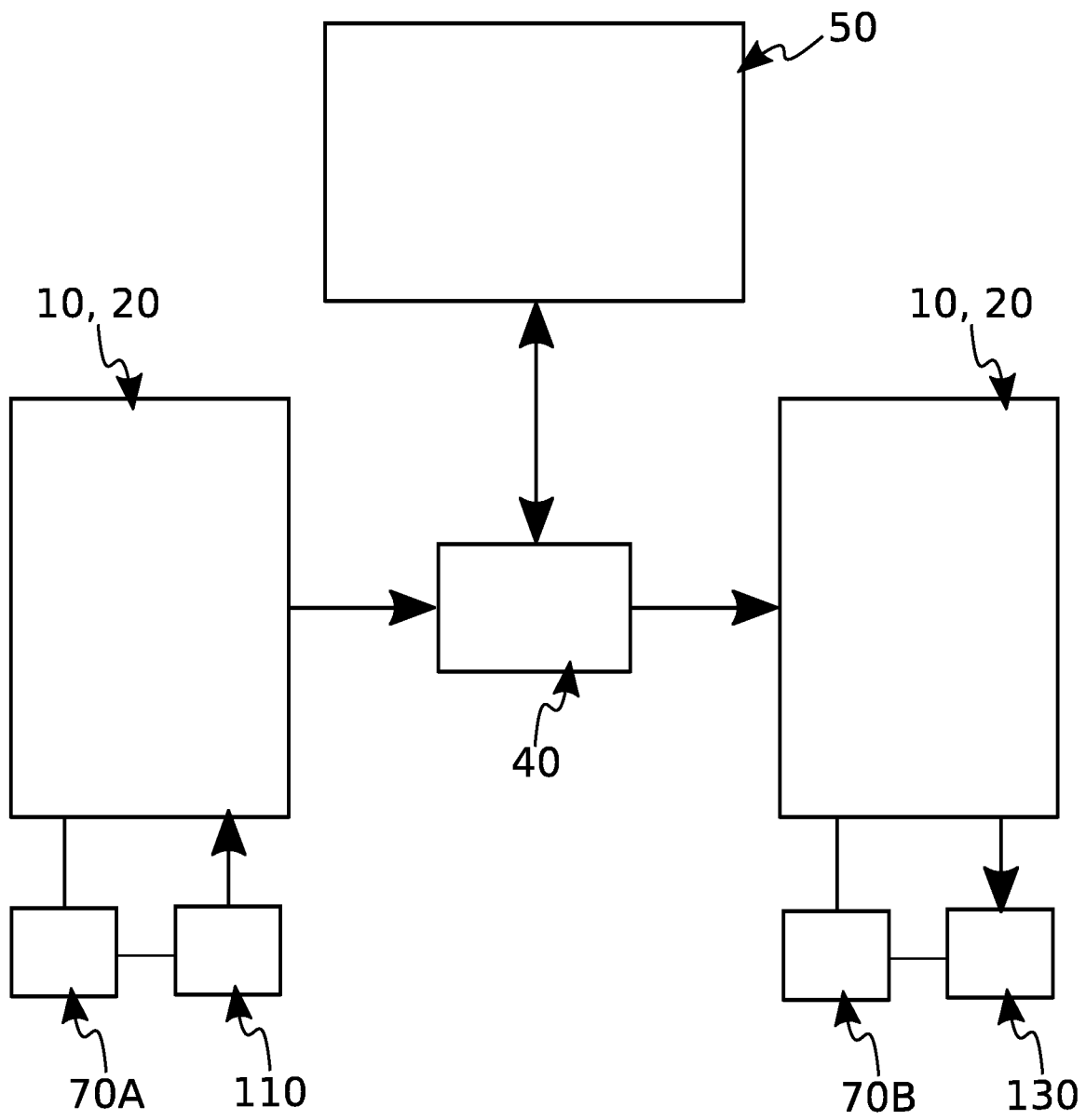
- obtain measurement related data (10, 20) configured according to a first code (70A);
- compare the obtained measurement related data (10, 20) with reference data (50);
- configure the measurement related data (10, 20) according to a second code (70B) based on the comparison and
- wherein the system is configured to carry out the method according to any of the preceding claims.

20. The communication system according to the preceding claim, further comprising a memory component configured to store the reference data (50),

a sending node (110) configured to generate the measurement related data (10, 20) according to the first code (70A),

a receiving node (130) configured to receive the measurement related data (10, 20) configured according to the second code (70B)

wherein the data processing system (40) is configured to carry out the method according to any of the preceding method claims 5 to 18 to configure the measurement related data (10, 20) generated by the sending node (110) according to the second code (70B).

**Fig. 1**

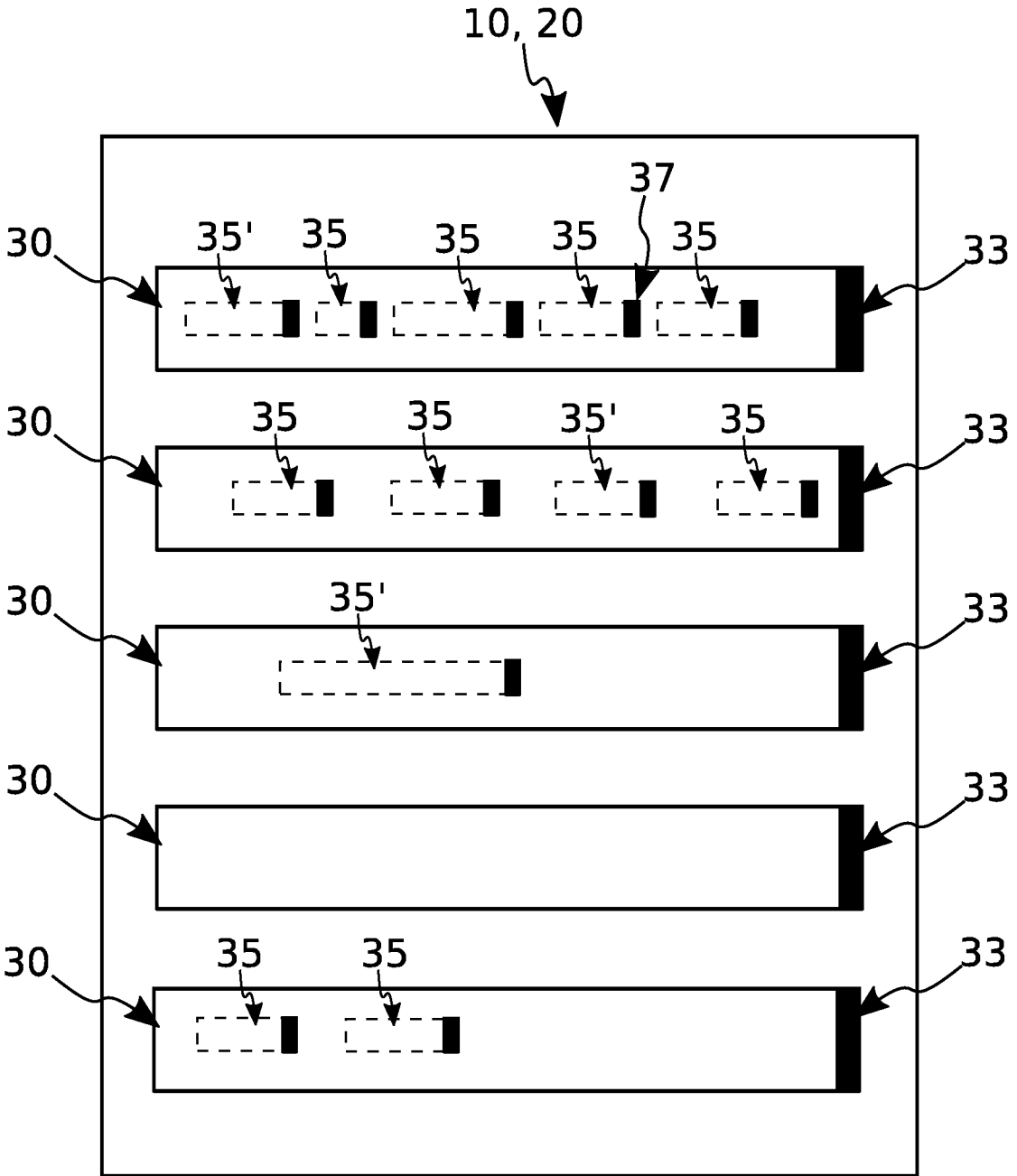


Fig. 2a

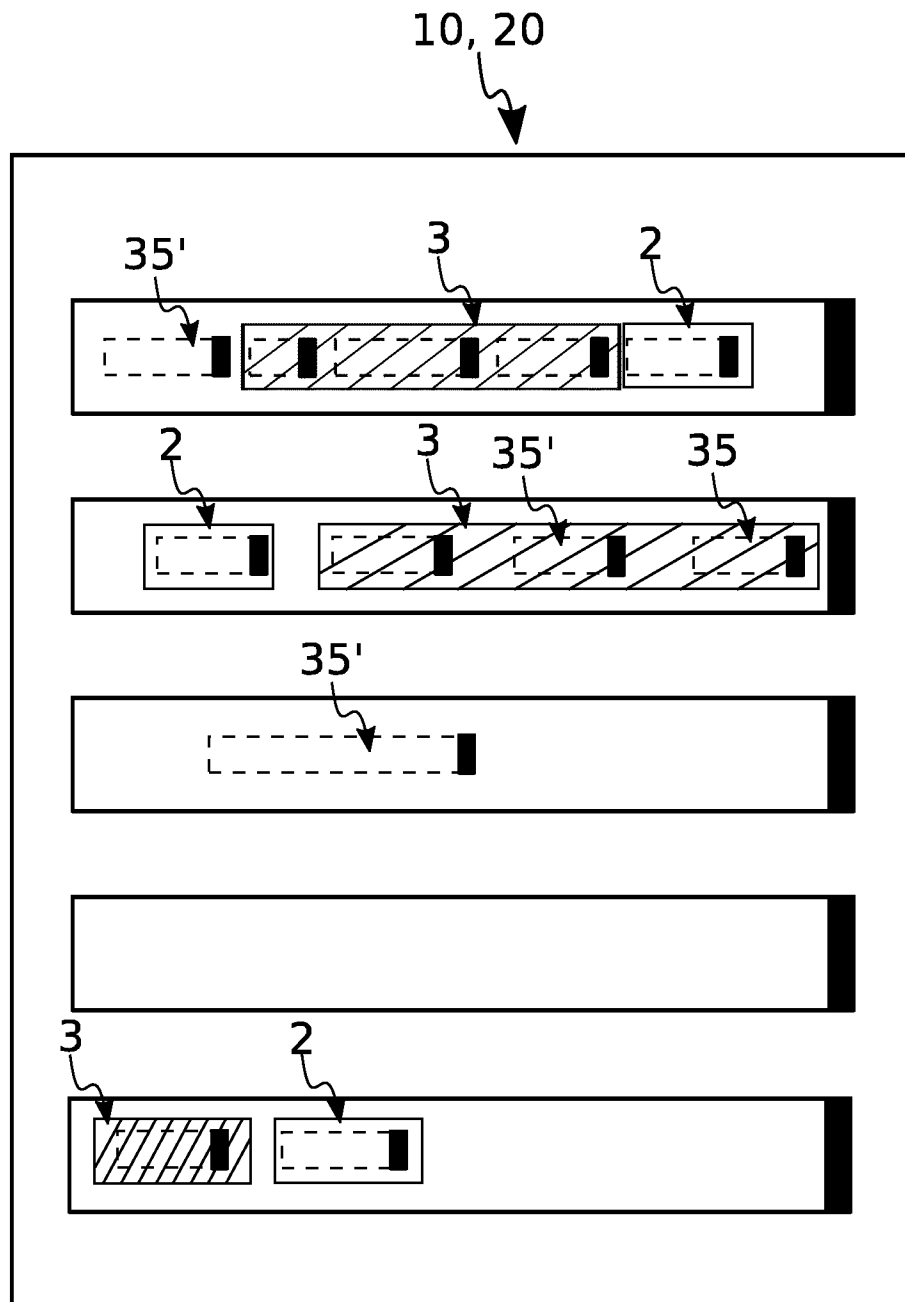
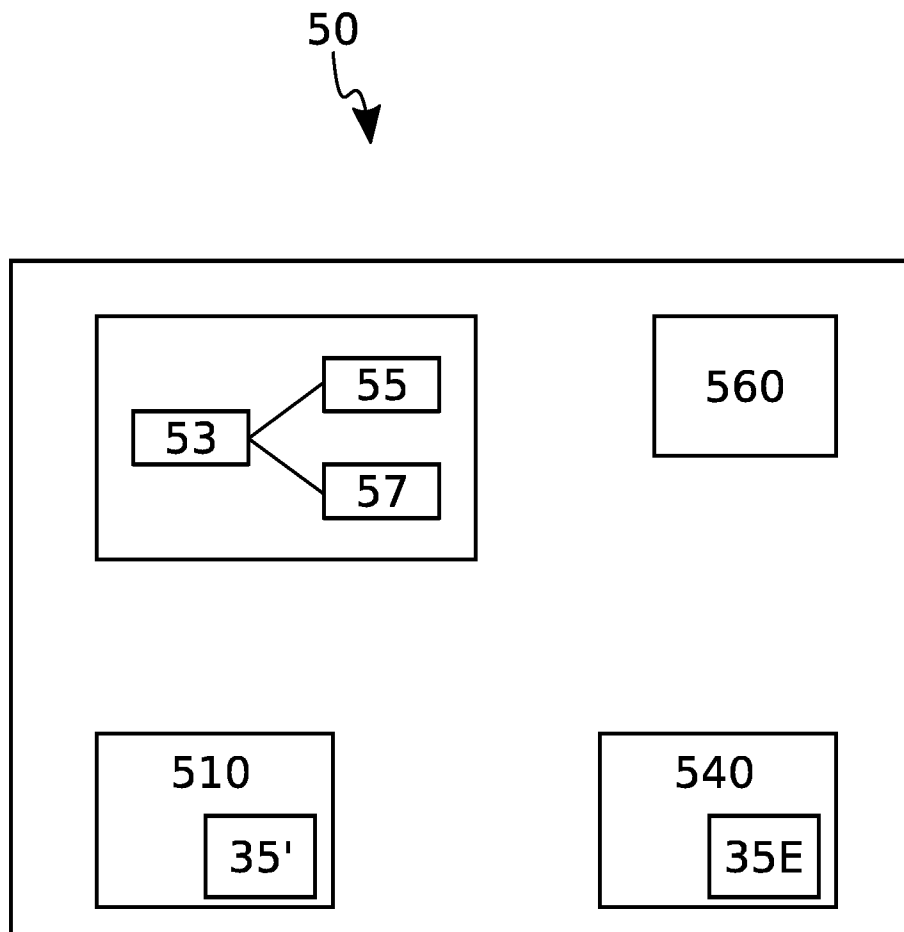
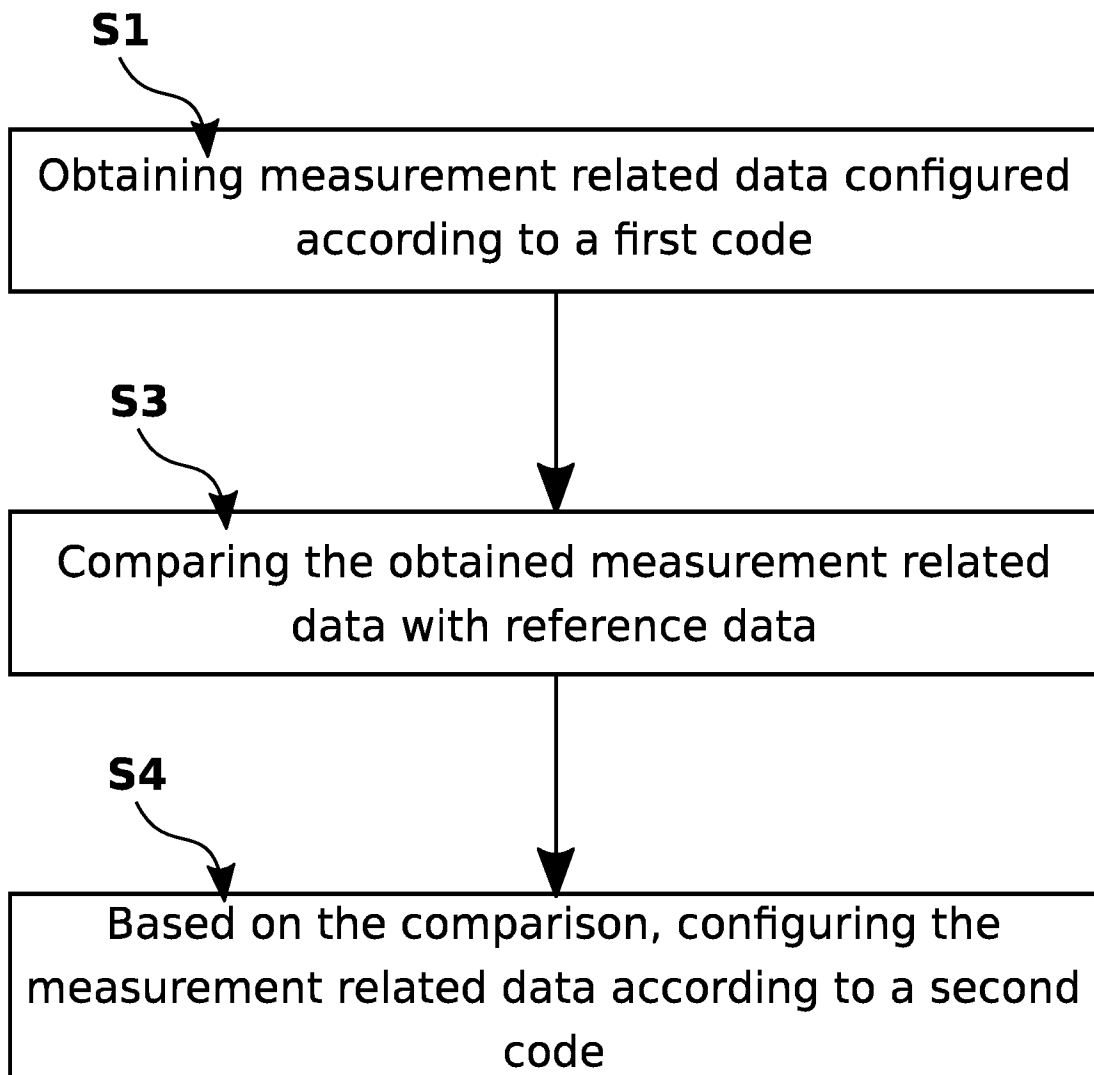
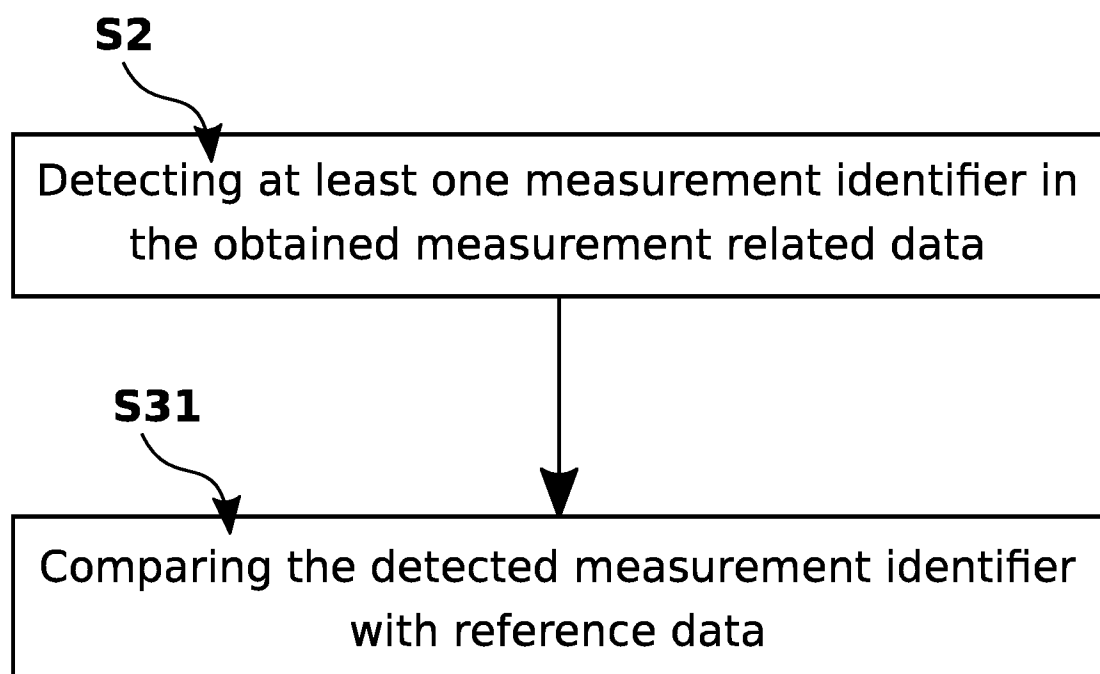
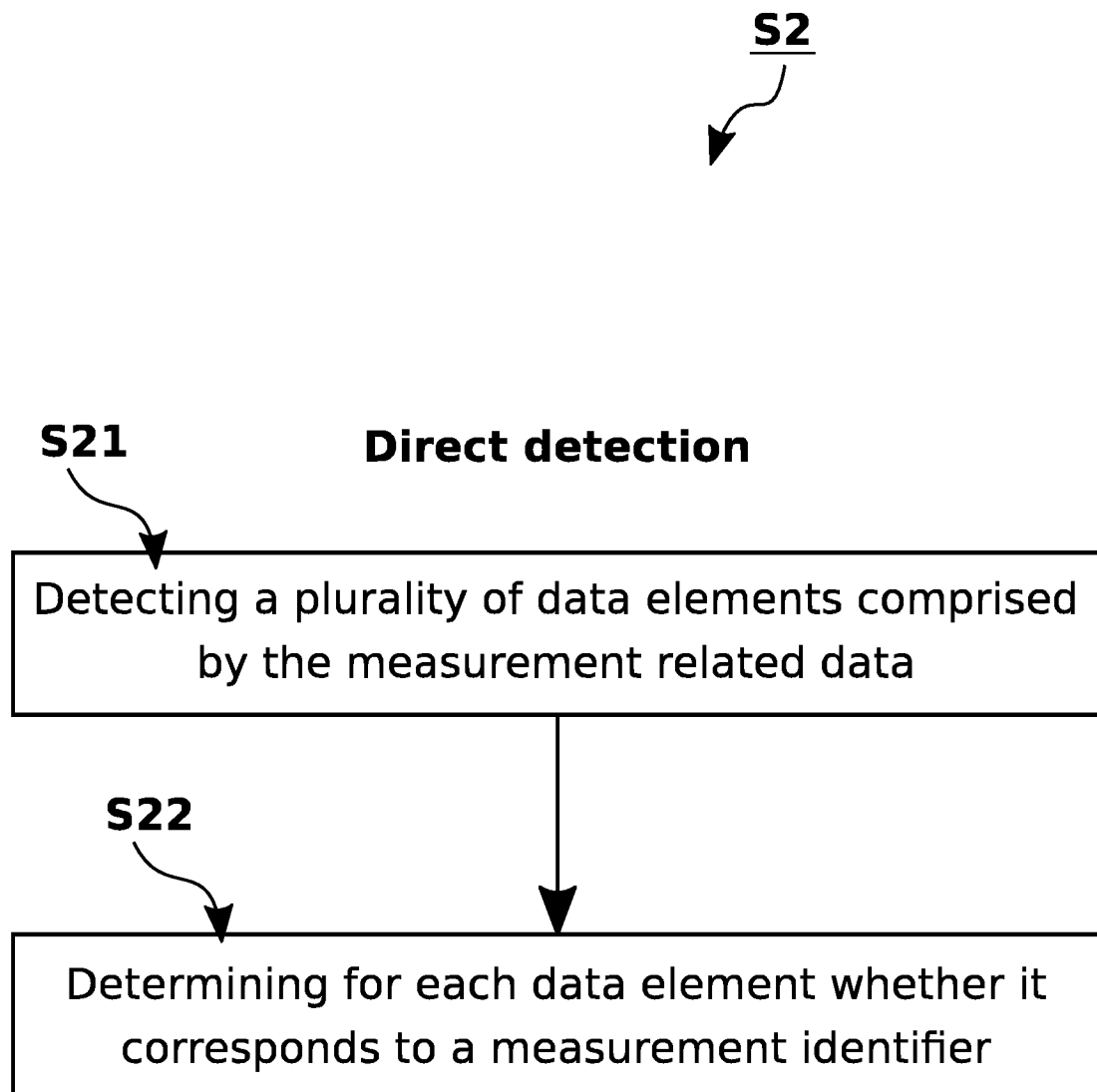


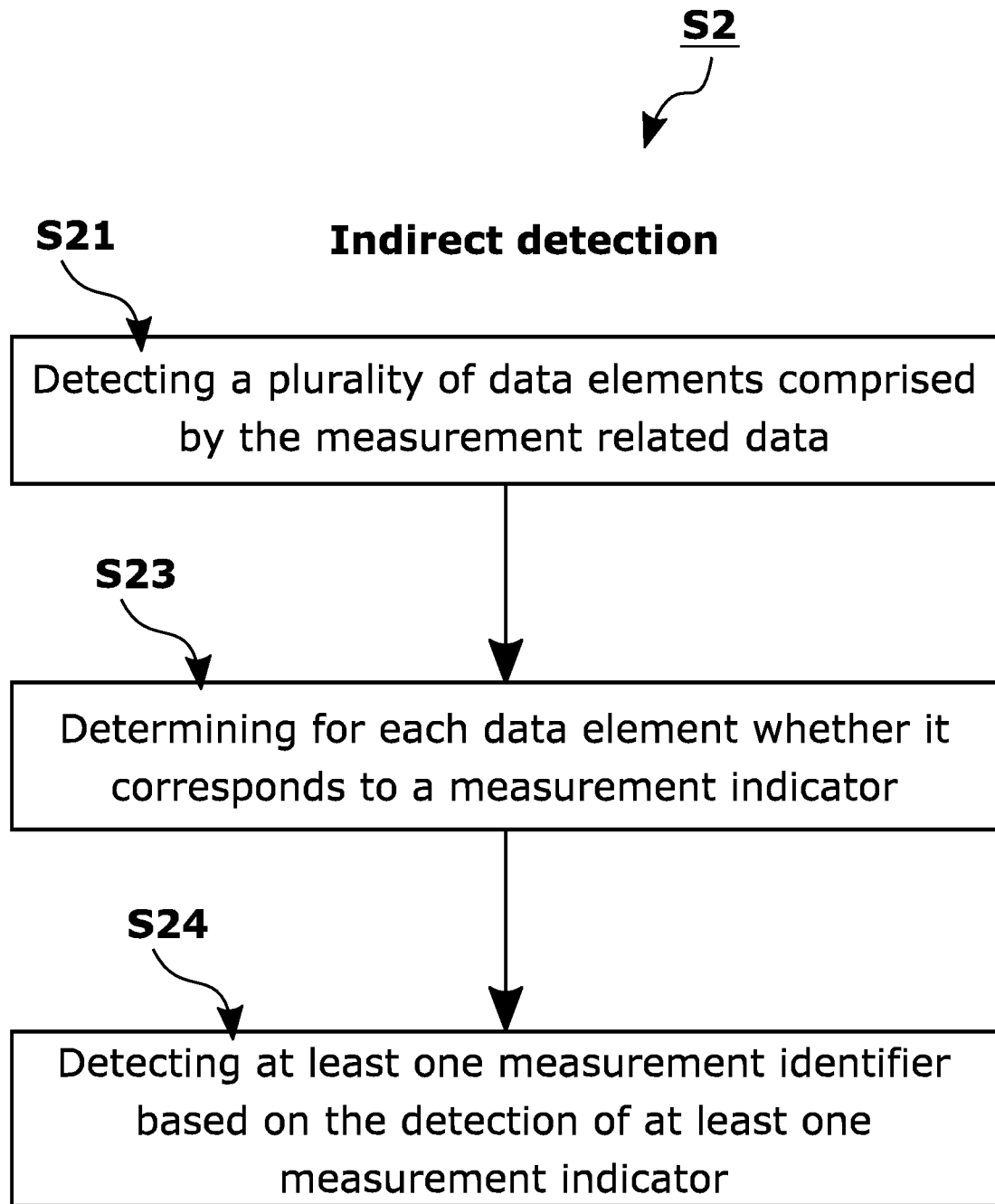
Fig. 2b

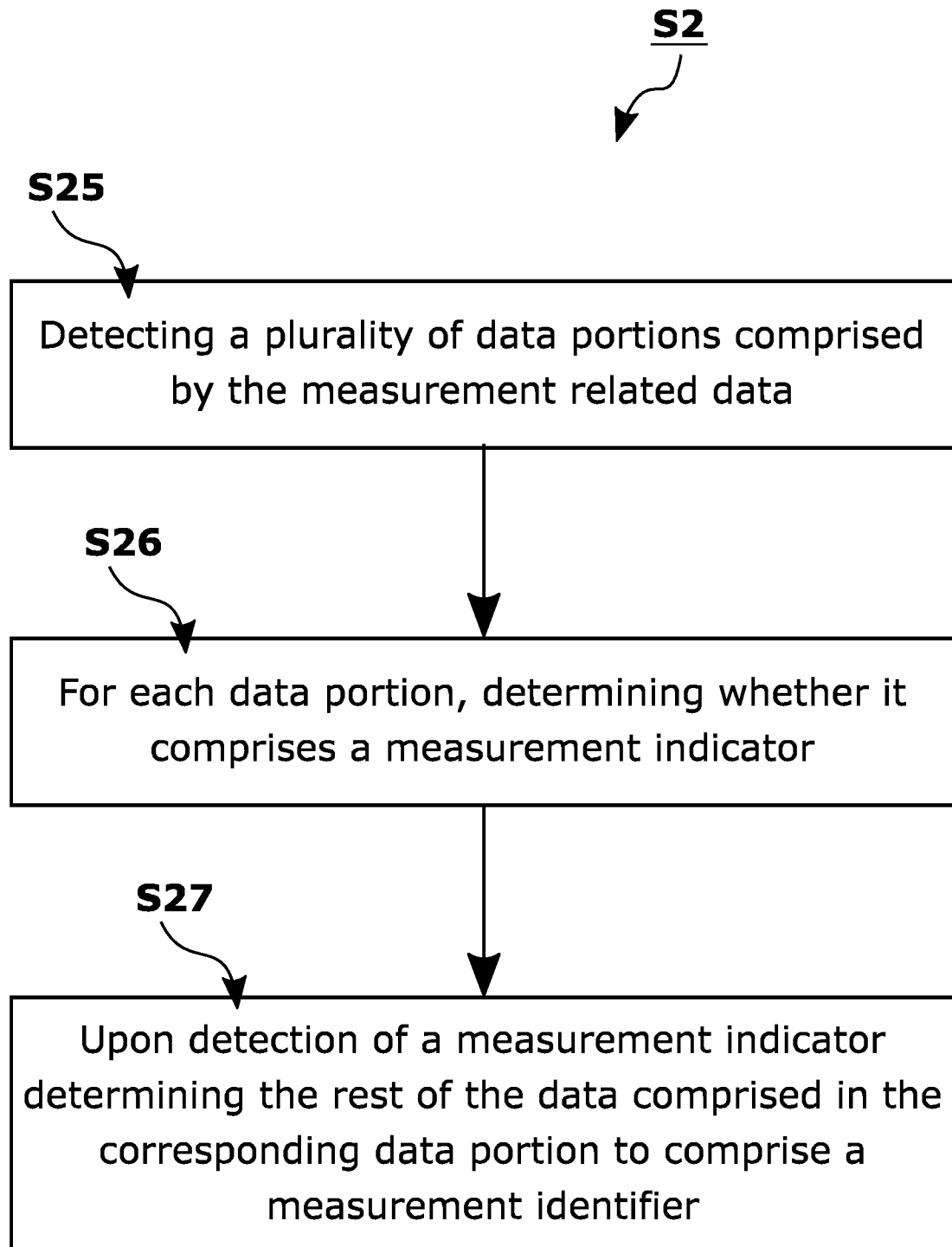
**Fig. 3**

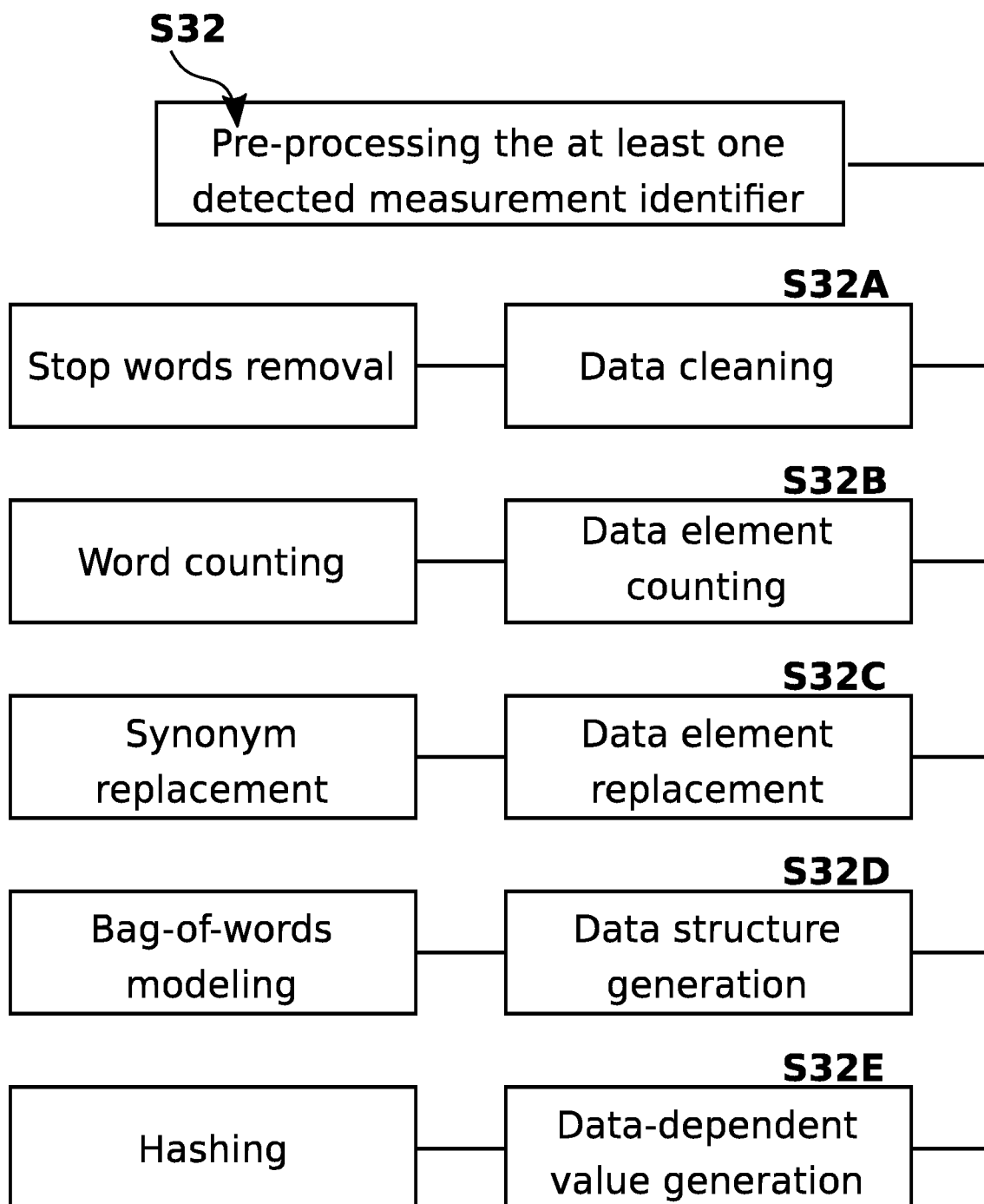
**Fig. 4a**

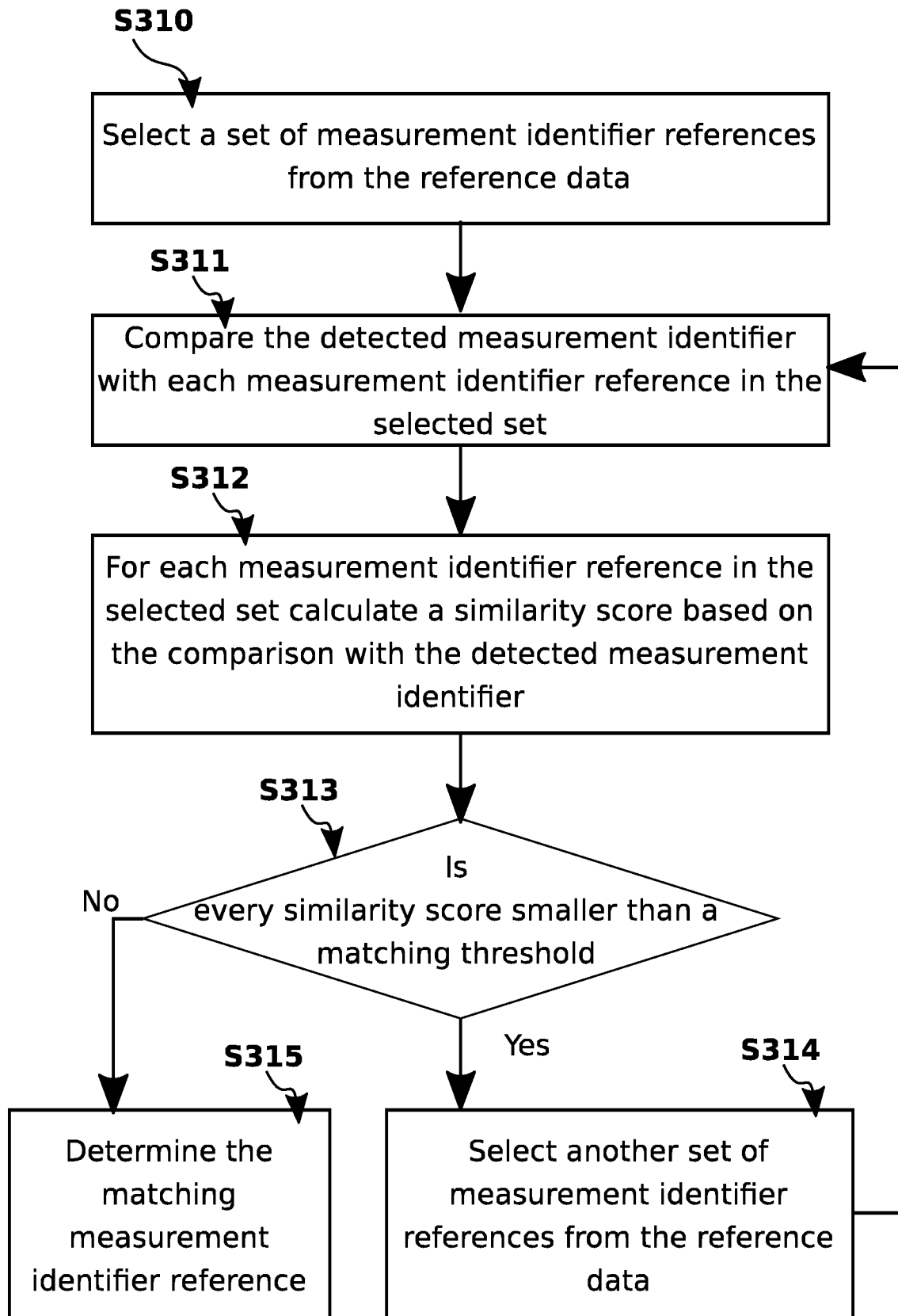
**Fig. 4b**

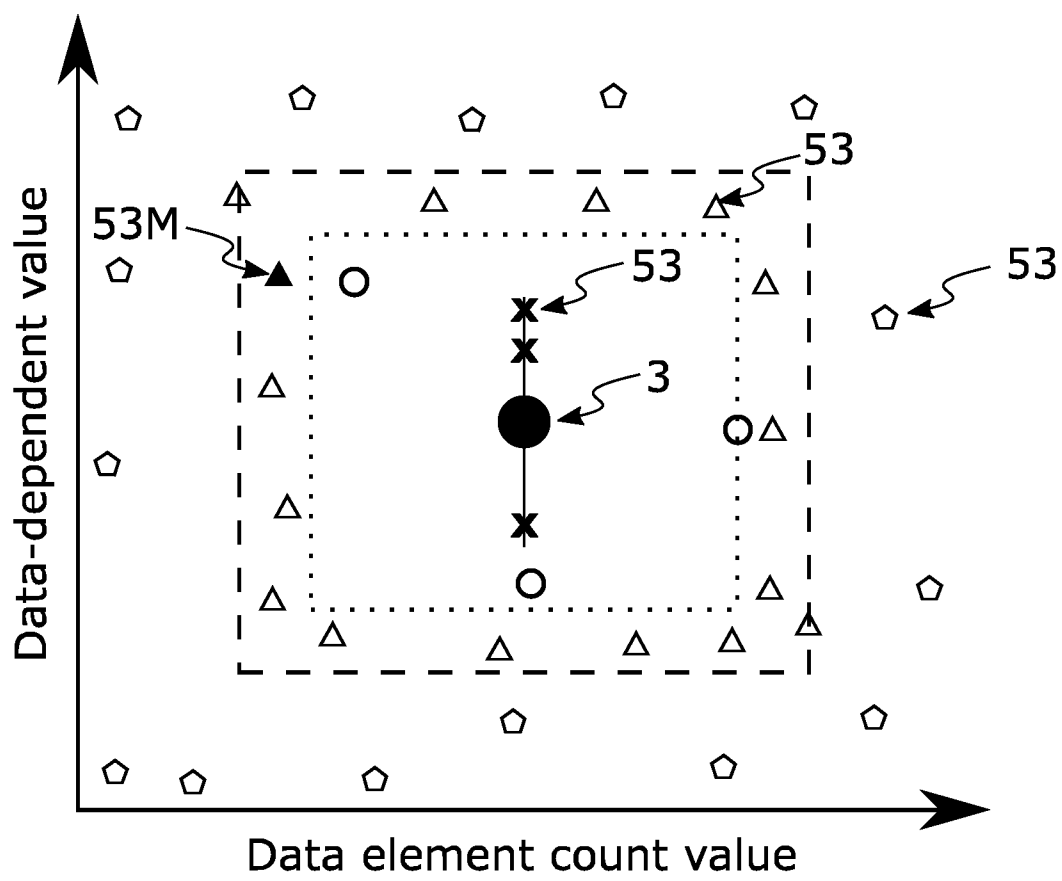
**Fig. 4c**

**Fig. 4d**

**Fig. 4e**

**Fig. 4f**

**Fig. 4g**

**Fig. 5**

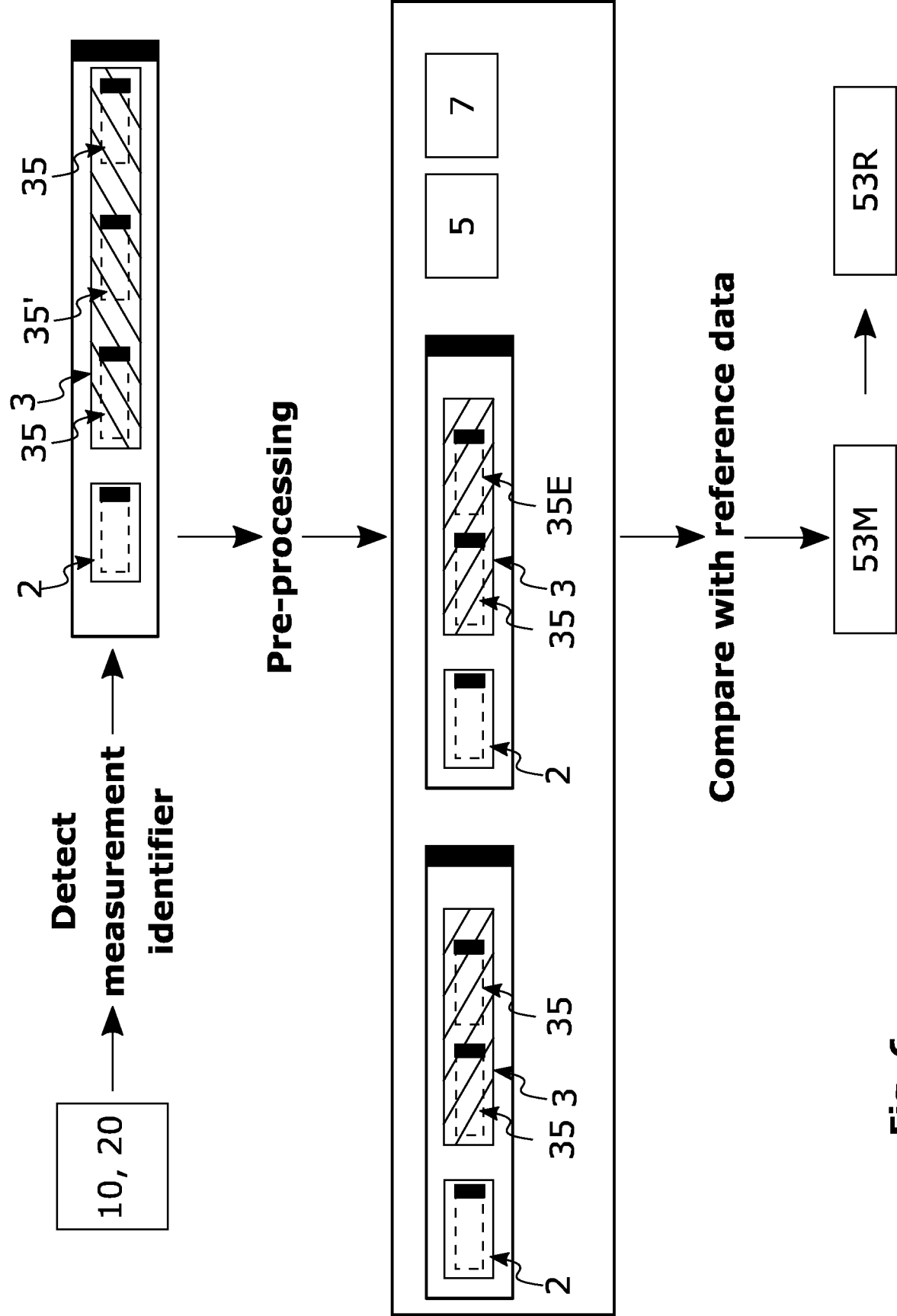


Fig. 6

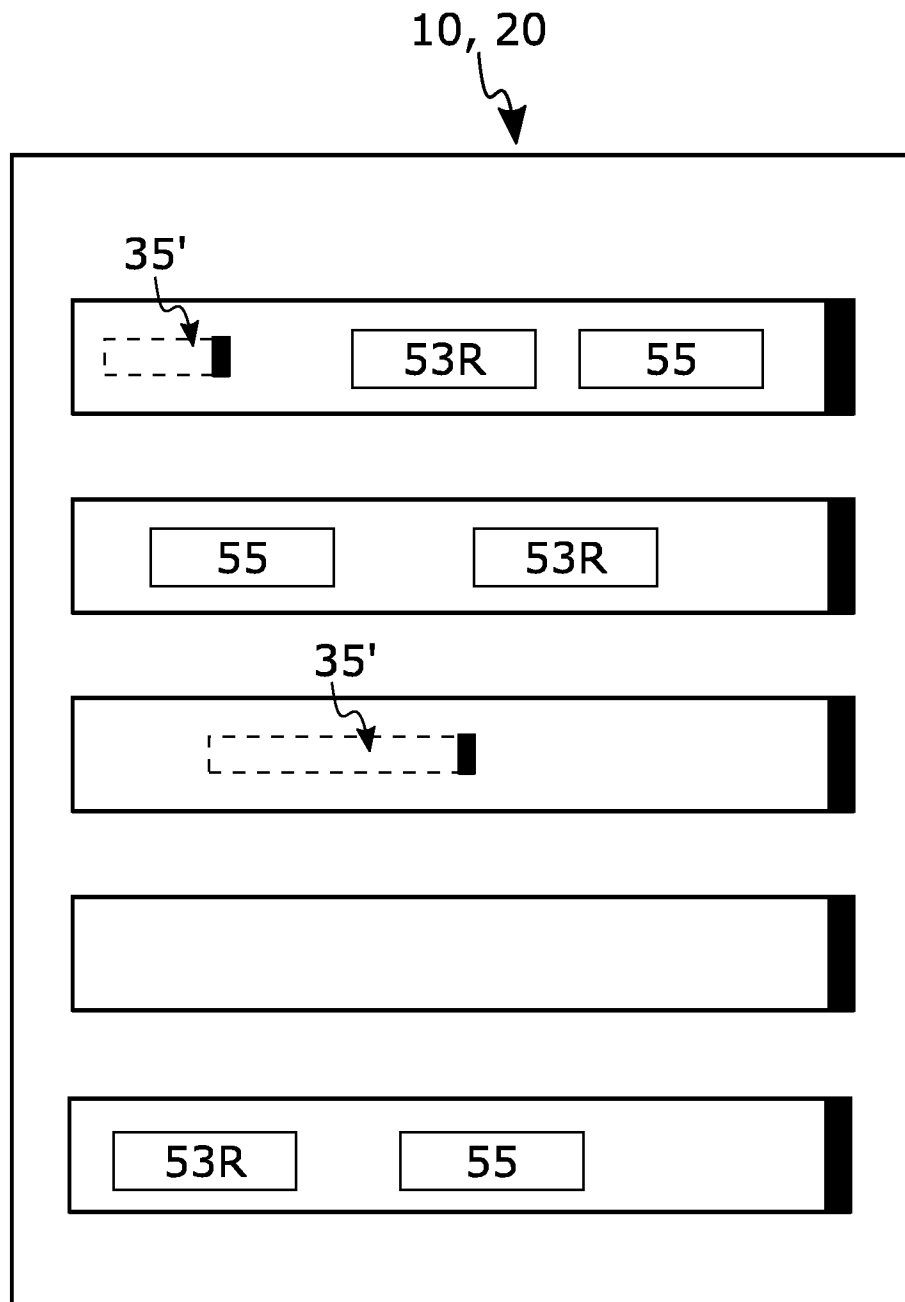


Fig. 7

50

53

55

57

70I

LOINC_NUM	LONG_COMMON_NAME	COMPONENT	PROPERTY	TIME_ASPECT	SYSTEM	SCALE_TYP	METHOD_TYP
14646-4	Cholesterol HDL [Moles/Volume] Serum/Plasma ; Numerique	Cholesterol.in HDL	SCnc	Pt	Ser/Plas	Qn	
2085-9	Cholesterol HDL [Masse/Volume] Serum/Plasma ; Numerique	Cholesterol HDL	MCnc	Pt	Ser/Plas	Qn	
49130-8	Cholesterol HDL [Masse/Volume] Serum/Plasma ; Numerique ; Electrophorese	Cholesterol HDL	MCnc	Pt	Ser/Plas	Qn	Electrophoresis
18263-4	Cholesterol HDL [Masse/Volume] Serum/Plasma ultracentrifuge ; Numerique	Cholesterol HDL	MCnc	Pt	Ser/Plas.ultracentrifugate	Qn	

Fig. 8a

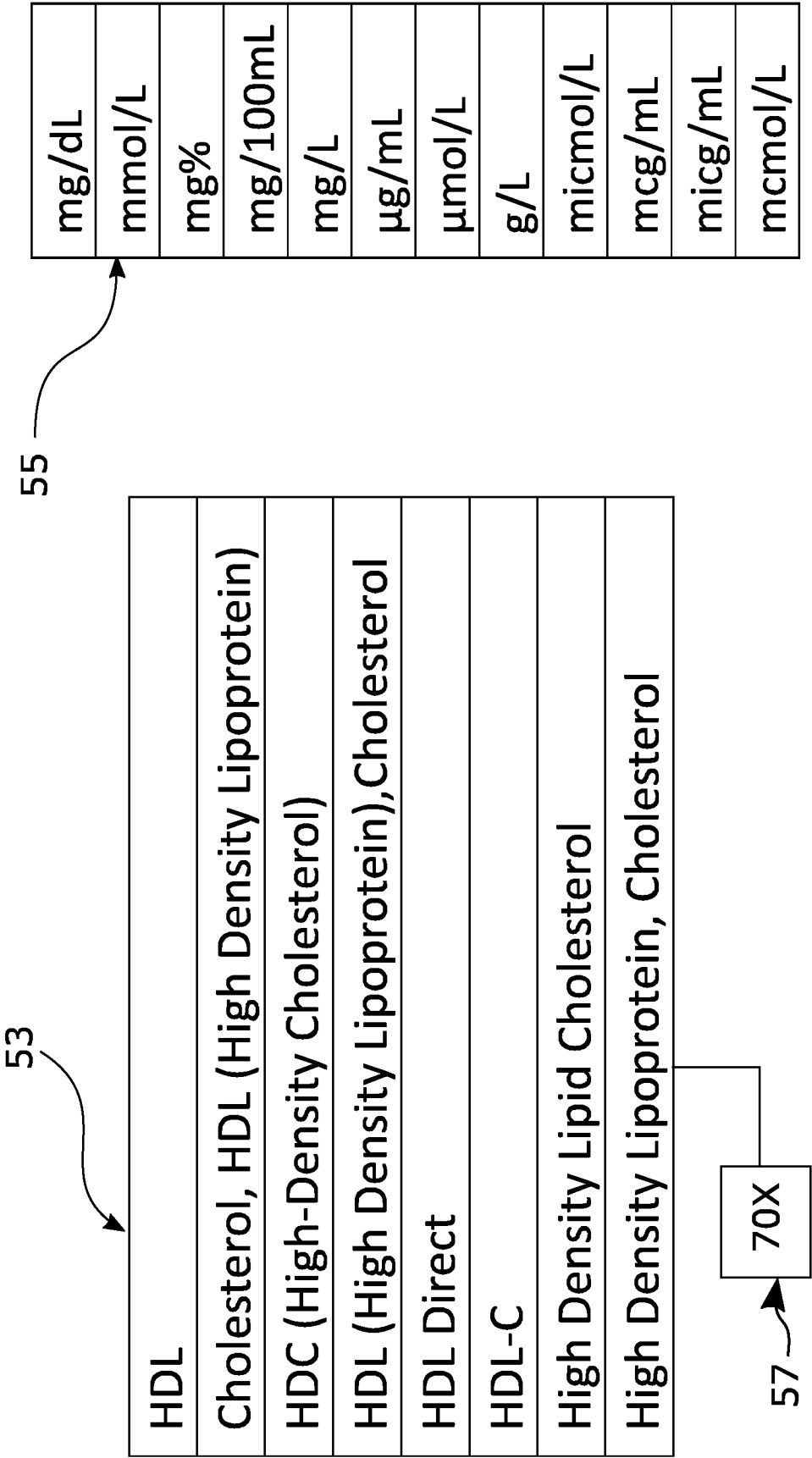


Fig. 8b

10, 20

3

2

2

2

2

Catalog Name	Unit	SYSTEM	METHOD_TYP	TIME_ASPCT	SCALE_TYP
HDL Chol.					
Alpha 1 Antitrypsin Phenotype, Serum					
Angiotensin Converting Enzyme, CSF					
B. Natriuretic Peptide					
Calcium, Ionized, Serum					
C-Reactive Protein, Quantitative					
Estradiol, Ultrasensitive, Serum or Plasma					
Estrogen, Total					

Fig. 8c

3

53M

55

55

55

55

55

9

Catalog Name	LOINC							Similarity
HDL Chol.	14646-4	Cholesterol in HDL	SCnc	Pt	Ser/Plas			95
HDL Chol.	2085-9	Cholesterol HDL	MCnc	Pt	Ser/Plas	Qn	Qn	95
HDL Chol.	49130-8	Cholesterol HDL	MCnc	Pt	Ser/Plas	Qn	Electrophoresis	95
HDL Chol.	18263-4	Cholesterol HDL	MCnc	Pt	Ser/Plas.ultracentrifugate	Qn	Qn	95

Fig. 8d

Biomarker in Report	Value	Unit in Report						
TOTAL LEUCOCYTES COUNT	5.08	X 10 ³ /μL						
Biomarker Abbreviation	Biomarker Name	Similarity	Biomarker Unit	Range	Min	Max		
WBC	Leucocyte Count	100	10e3/μL	From-To	4.0	10.0		

3 → TOTAL LEUCOCYTES COUNT

2 → Value

2 → Unit in Report

53M, 53R → Leucocyte Count

9 → Similarity

55 → Biomarker Unit

2 → Range From-To

2 → Min

2 → Max

Fig- 9b

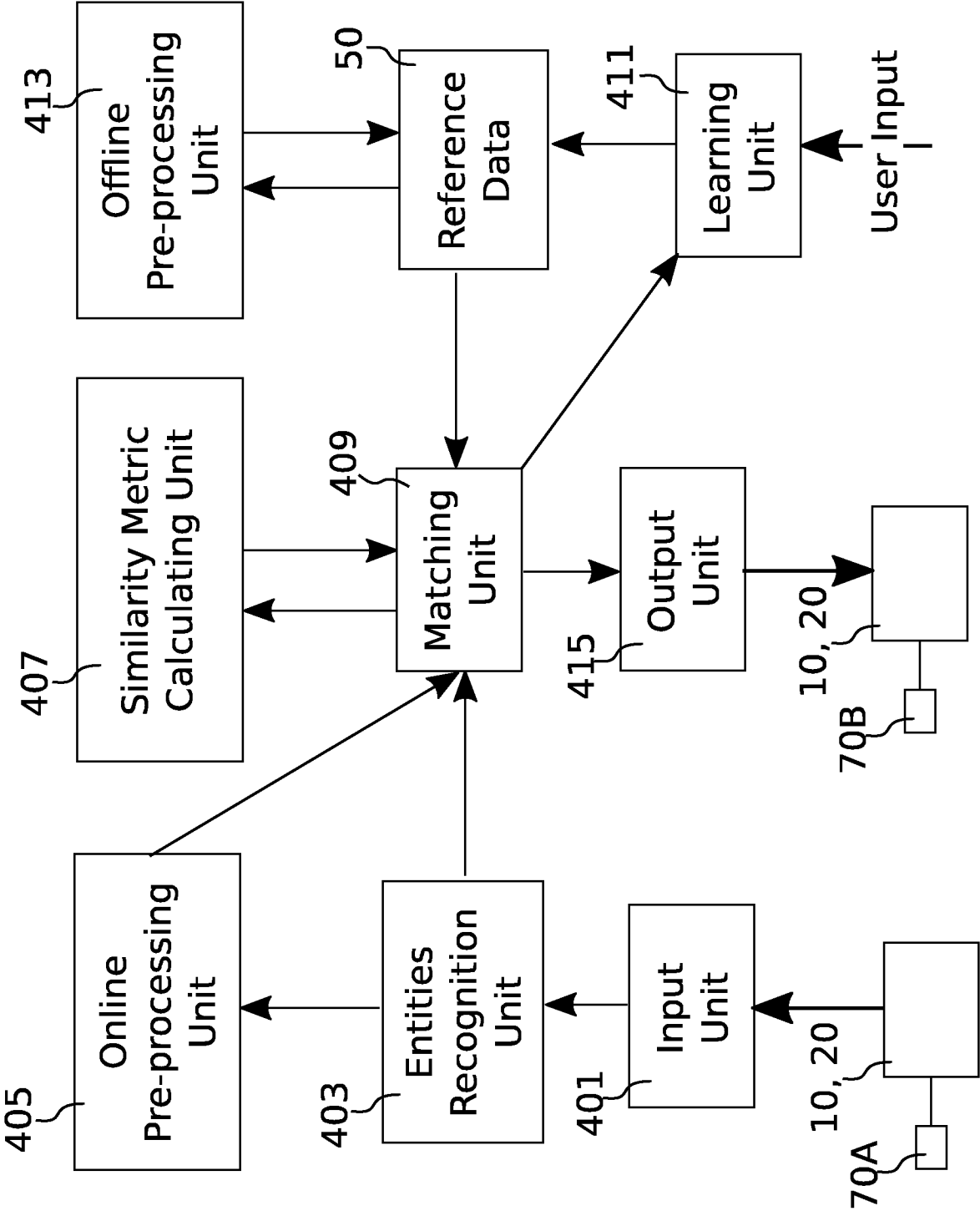


Fig. 10

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2021/051997

A. CLASSIFICATION OF SUBJECT MATTER

INV. G16H10/40
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G16H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2008/270438 A1 (ARONSON SAMUEL J [US] ET AL) 30 October 2008 (2008-10-30) The whole document, in particular: Paragraphs [0050] - [0053], [0061], [0076] - [0085], [0089] - [0091], [0096]; Figures 2, 3, 4, 6, 7. -----	1-20
X	US 5 392 209 A (EASON REGINALD [US] ET AL) 21 February 1995 (1995-02-21) The whole document, in particular: Claims 12 - 15; Columns 10 - 14. -----	1-20
X	US 2011/257998 A1 (CINQUALBRE JACQUES [FR] ET AL) 20 October 2011 (2011-10-20) The whole document in particular: Paragraphs [0012], [0023] - [0026], [0033] - [0042], [0062], [0139] - [0141]; Figures 1, 2, 4. ----- -/--	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

27 April 2021

Date of mailing of the international search report

07/05/2021

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Nagele, Stefan

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2021/051997

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2002/198739 A1 (LAU LEE MIN [US] ET AL) 26 December 2002 (2002-12-26) The whole document, in particular: Paragraphs [0006], [0015], [0040] - [0053], [0056] - [0059], [0062], [0063]. -----	1-20
X	US 2006/235881 A1 (MASARIE FRED E [US] ET AL) 19 October 2006 (2006-10-19) The whole document, in particular: Paragraphs [0012], [0014] - [0017], [0022] - [0026], [0046] - [0054]. -----	1-20
X	US 2020/020441 A1 (VIJAY SANTOSH V [US]) 16 January 2020 (2020-01-16) The whole document, in particular: Paragraphs [0001], [0002], [0013] - [0019], [0033] - [0056]; Figures 1, 2. -----	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2021/051997

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008270438 A1	30-10-2008	AU 2008216177 A1 CA 2678128 A1 EP 2115650 A2 JP 2010521722 A KR 20090110385 A US 2008270438 A1 US 2011119309 A1 WO 2008101114 A2	21-08-2008 21-08-2008 11-11-2009 24-06-2010 21-10-2009 30-10-2008 19-05-2011 21-08-2008
US 5392209 A	21-02-1995	AU 5870494 A CA 2148266 A1 EP 0674782 A1 JP H08504988 A US 5392209 A WO 9415270 A1	19-07-1994 07-07-1994 04-10-1995 28-05-1996 21-02-1995 07-07-1994
US 2011257998 A1	20-10-2011	FR 2953961 A1 US 2011257998 A1	17-06-2011 20-10-2011
US 2002198739 A1	26-12-2002	NONE	
US 2006235881 A1	19-10-2006	CN 101198952 A EP 1875382 A2 JP 2008537227 A KR 20070120152 A US 2006235881 A1 WO 2006113298 A2	11-06-2008 09-01-2008 11-09-2008 21-12-2007 19-10-2006 26-10-2006
US 2020020441 A1	16-01-2020	US 2020020441 A1 WO 2020018490 A1	16-01-2020 23-01-2020