

DATA CONVERSION AND DISTRIBUTION SYSTEMS

TECHNICAL FIELD

[0001] The present disclosure relates generally towards improving electronic data conversion and distribution, and, in particular to systems and methods for electronic data conversion and distribution of electronic data sensitivities and projections where electronic data is sparse, whether from high volume data sources and/or differently formatted electronic data sources.

BACKGROUND

[0002] Problems exist in the field of electronic data conversion and distribution. Users of data classes with sparse electronic data often seek additional data and information in order to analyze or otherwise utilize these data classes. One utilization of electronic data is in the creation of data projections (or other statistical analyses / applications) for those data classes having sparse electronic data (e.g., limited historical data). Since the electronic data is sparse, it may be a challenge to obtain the additional electronic data and information needed, at desired time(s) and/or in desired data types and volumes, to generate accurate data projections. Indeed, accurate projections (and other forms of statistical analysis) typically require a large amount of historic electronic data and/or information for analysis. In the absence of such data and information, conventional projections (based on the sparse data and information) are often very inaccurate and unreliable. Accordingly, there is a need for improved data conversion and distribution systems which are able to generate accurate projections and yield other data analysis results that are accurate and timely, even if the data being projected is sparse.

SUMMARY

[0003] The present disclosure is related to data conversion and distribution systems which are able to process and utilize any amount of data, received at different volumes, frequencies, and/or formats, from any number of different data sources in order to generate data that is usable for creating accurate data sensitivities, projections and/or yielding other statistical analyses associated with a data class having sparse data, all in a timely manner.

[0004] Aspects of the present disclosure include systems, methods and non-transitory computer-readable storage media specially configured for data conversion and distribution. The systems, methods, and non-transitory computer readable media may further include a data subscription unit and a virtual machine. The data subscription unit may have at least one data interface communicatively coupled to a plurality of data source devices and may be configured

to obtain data having a plurality of data formats from the plurality of data source devices. The data subscription unit may also be configured to transmit the data having the plurality of data formats via secure communication over a network. The virtual machine of the present disclosure may include one or more servers, a non-transitory memory, and/or one or more processors including machine readable instructions. The virtual machine may be communicatively coupled to the data subscription unit. The virtual machine may further include a data receiver module, a data unification module, a data conversion module, and/or a data transmission module. The data receiver module of the virtual machine may be configured to receive the data having the plurality of data formats from the data subscription unit via the secure communication over the network. The data unification module of the virtual machine may be configured to reformat and/or aggregate the data from the data subscription unit to generate unified data responsive to receiving, at the receiver module, the data having the plurality of data formats. The data conversion module may be configured to run the unified data through one or more statistical algorithms in order to generate at least one of data sensitivities and projected data based on the unified data. The data transmission module may be configured to transmit the at least one of the data sensitivities and the projected data to a data distribution device via one or more secure communications over a network.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] FIG. 1A is a functional block diagram of an embodiment of a data conversion and distribution system in accordance with the present disclosure.

[0006] FIG. 1B is a flowchart of an example method for data conversion and distribution in accordance with the present disclosure.

[0007] FIG. 2 is a functional block diagram of a data subscription unit in accordance with an embodiment of a data conversion and distribution system of the present disclosure.

[0008] FIG. 3 is a functional block diagram of a virtual machine in accordance with an embodiment of a data conversion and distribution system of the present disclosure.

[0009] FIG. 4 is a flowchart of an example statistical algorithm for generating data sensitivities and/or projected data in accordance with an embodiment of a data conversion and distribution system of the present disclosure.

[0010] FIG. 5 is a functional block diagram of a data distribution device in accordance with an embodiment of a data conversion and distribution system of the present disclosure.

[0011] FIG. 6 is a functional block diagram of a remote user device in accordance with an embodiment of a data conversion and distribution system of the present disclosure.

[0012] FIG. 7 is a schematic representation of a graphical user interface used in connection with an embodiment of the present disclosure.

[0013] FIG. 8 is a schematic representation of a graphical user interface used in connection with an embodiment of the present disclosure.

DETAILED DESCRIPTION

[0014] Aspects of the present disclosure relate to systems, methods and non-transitory computer-readable storage media for data conversion and distribution.

[0015] An example data conversion and distribution system of the present disclosure may include a data subscription unit and a virtual machine. The data subscription unit may have at least one data interface communicatively coupled to a plurality of data source devices and may be configured to obtain data having a plurality of data formats from the plurality of different data source devices. The data subscription unit may also be configured to transmit the data having the plurality of data formats via secure communication over a network. The virtual machine of the system may include one or more servers, a non-transitory memory, and one or more processors including machine readable instructions. The virtual machine may be communicatively coupled to the data subscription unit. The virtual machine may also include a data receiver module, a data unification module, a data conversion module, and/or a data transmission module. The data receiver module of the virtual machine may be configured to receive the data having the plurality of data formats from the data subscription unit via the secure communication over the network. The data unification module of the virtual machine may be configured to reformat and aggregate the data (having the plurality of data formats) from the data subscription unit, to generate unified data responsive to receiving, at the receiver module, the unified data having a standardized data format. The data conversion module may be configured to run the unified data through one or more statistical algorithms in order to generate at least one

of data sensitivities and projected data for a data class that is not necessarily directly related to the data received from the plurality of data sources. In other words, the unified data, which originates from a plurality of data sources *other than* that of the data class and which may be indirectly or tangentially related to the data class, may be used to generate data sensitivities, data projections and/or other statistical information representative of the data class. The data transmission module may be configured to transmit the at least one of the data sensitivities and the projected data to a data distribution device via one or more secure communications over a network.

[0016] In one embodiment, the data distribution device further includes a non-transitory memory and at least one data distribution interface. The non-transitory memory may be configured to store the at least one of the data sensitivities and the projected data. One or more of the data distribution interfaces may be configured to provide secure communications with at least one of one or more remote user devices.

[0017] In one embodiment, a remote user device may include a non-transitory memory, one or more processors including machine readable instructions, a data distribution receiver interface communicatively coupled to the data distribution device, a user information interface, a market data source interface, and/or a user display interface. One or more of the remote user devices may be further configured to receive the data sensitivities and/or the projected data from the data distribution device via the data distribution receiver interface, receive user input data via the user information interface, receive current market data via the market data source interface, generate supplementary projected data via one or more processors and/or display at least a portion of the projected data and the supplementary projected data on a user display interface. The supplementary projected data may be based on the received data sensitivities, projected data, user input data, and/or current market data.

[0018] An exemplary embodiment of a data conversion and distribution system 100 is illustrated in FIG. 1A. As depicted, the data conversion and distribution system 100 may include a data subscription unit 101, a virtual machine 103, and a data distribution device 105. The data subscription unit 101, the virtual machine 103 and the data distribution device 105 may be communicatively coupled via a network 108. Alternatively or additionally, the data subscription unit 101 may be directly coupled to the virtual machine 103, and/or the virtual machine 103 may

be directly coupled to the data distribution device 105, without the use of a network. The data conversion and distribution system 100 may further include one or more remote user devices 107. In one example, each of the remote user devices 107 may be used by participants including for example, data managers, data analysts, regulatory compliance teams, and the like. Although system 100 is described in some examples below with respect to data classes associated with electronic instrument data, system 100 may be used with any electronic data classes associated with any type of electronic data, including those having sparse data. The data subscription unit 101 may have at least one data interface (e.g., data interface 201 shown in FIG. 2) communicatively coupled to one or more data source devices 109. Although the description and drawings herein describe the data conversion and distribution system 100 and its surrounding environment as having one or more data source devices 109 (Data Source Device 1 – Data Source Device N) and one or more remote user devices 107 (Remote User Device 1 – Remote User Device N), in some examples, there may be any combination of data source devices 109 and/or remote user devices 107, including for example, a single data source device 109 and a single remote user device 107, or a single data source device 109 and no remote user devices 107. One or more of the data source devices 109, data subscription unit 101, virtual machine 103, data distribution device 105, and remote user devices 107 may include one or more computing devices including a non-transitory memory component storing computer-readable instructions executable by a processing device to perform the functions described herein.

[0019] The data source devices 109 may be communicatively coupled to the data subscription unit 101 via a network 110. The data distribution device 105 may be communicatively coupled to the remote user devices 107 via a network 106. In some embodiments, the networks 110 and 106 may include two or more separate networks to provide additional security to the remote user devices 107 by preventing direct communication between the remote user devices 107 and the data source devices 109. Alternatively, the networks 110, 106 may be linked and/or a single large network. The networks 110, 106 (as well as network 108) may include, for example, a private network (e.g., a local area network (LAN), a wide area network (WAN), intranet, etc.) and/or a public network (e.g., the internet). Networks 110 and/or 106 may be separate from or connected to network 108.

[0020] FIG. 1B is a flowchart of an example method corresponding to the data conversion and distribution system 100 of FIG. 1A (also described with respect to FIGS. 2, 3, 5 and 6). As illustrated in FIG. 1A, a method for data conversion and distribution may include, at step 121, obtaining data having a plurality of data formats from the data source devices 109. The data source devices 109 may include data and information directly, indirectly and/or tangentially related to the data class. The data source devices 109 may be selected based on their perceived relevance to the data class and/or usefulness in statistical calculations (e.g., generating data projections) for the data class having limited or sparse data. In one embodiment, the data source devices 109 may be selected by way of subscription preferences designated by a remote user device 107 and/or by an operator of the data conversion and distribution system 100 itself. Additionally, the data obtained from the data source devices 109 may be ‘cleansed’ (which may involve analyzing, filtering and/or other operations discussed in further detail below) to ensure that only pertinent data and information is used in the statistical calculations, thereby improving the accuracy of any resulting calculations while at the same time reducing the amount of data and information that must be modeled (i.e., run through statistical algorithms that execute the statistical calculations). The data may be obtained, for example, via data interface 201 of the data subscription unit 101. Step 121 is described further below with respect to FIG. 2.

[0021] In step 123, the data having the plurality of data formats may be transmitted, for example, by data transmitter 207 of the data subscription unit 101, to the virtual machine 103 via network 108. Step 123 is discussed further below with respect to FIG. 2.

[0022] At step 125, a data receiver module 307 of the virtual machine 103 may receive the data having the plurality of data formats from the data subscription unit 101. At step 127, the data received from the data subscription unit 101 may be reformatted and aggregated (discussed below), for example, by data unification module 309 of virtual machine 103, to form unified data. Optionally, the data unification module 309 of the virtual machine 103 may also unpack and/or cleanse (discussed below) the data prior to forming unified data. Steps 125 and 127 are discussed further below with respect to FIG. 3.

[0023] At step 129, the data conversion module 311 of the virtual machine 103 may run the unified data through any number of algorithms (e.g., statistical algorithms) to generate data sensitivities, data projections, and/or any other desired statistical analyses information. Step 129

is discussed further below with respect to FIG. 3. An example algorithm of step 129 is also described further below with respect to FIG. 4.

[0024] At step 131, the generated data sensitivities, projected data and/or other statistical analyses information may be transmitted, for example, via the data transmission module 315 of the virtual machine 103, to a data distribution device 105. The transmission may be performed using one or more secure communications over the network 108. Step 131 is described further below with respect to FIG. 5.

[0025] At step 133, the data distribution device 105 may transmit at least a portion of the generated data sensitivities, projected data and/or other statistical analyses information to one or more remote user devices 107, for example, in response to a request received from among the remote user devices 107. Step 133 is described further below with respect to FIGS. 5 and 6.

[0026] The data source devices 109 of FIG. 1A may include additional electronic data and/or other information useful for supplementing and/or making statistical determinations for sparse electronic data sets. In general, the electronic data, and/or information may include suitable real-time data and/or archived data which may be related to a data class having sparse data and which may be useful for determining data sensitivities, data projections and/or statistical analyses information for the data class. In one example, the data source devices 109 of FIG. 1A may include internal and external data sources which may provide real-time and archived data. Internal data sources may include data sources that are a part of the particular entity seeking to supplement and/or generate statistical information for a data class that pertains to that particular entity; whereas external data sources may sources of data and information other than the entity that is seeking to supplement and/or generate the statistical information. For example, in one type of organization, the data source devices 109 may include internal data related to sales, purchases, orders, and transactions. The data sources may also include data aggregators. Data aggregators may store information and data related to multiple data classes. The data aggregators may themselves obtain the data and information from a plurality of other internal and/or external data sources. In some examples, the data sources may include information regarding current activity data, reference data and security information (all of which may vary by industry). In some examples, data sources of data source devices 109 may include news and media outlets, exchanges, regulators, and the like. Data source devices 109 may contain

information related to domestic and foreign products and/or services. In one embodiment, the data source devices 109 may contain information regarding quotes counts, trade counts, and trade volume.

[0027] Each of the data source devices 109 may produce one or more electronic data files. The electronic data files may include additional data and information pertinent to sparse electronic data. The additional data and information may be useful for generating data sensitivities, projections for sparse electronic data and/or statistical analyses information. In one example, the electronic data files may include data related to current activity, reference data, and security information. In another example, the electronic data files may include data related to pricing, market depth, dealer quotes, transactions, aggregate statistics, a quantity of products / instruments, a total par amount, advances, declines, highs and lows, and/or the like. Notably, any type of data may be included in the data files, depending on the particular industry and/or implementation of the data conversion and distribution system of the present disclosure. In one embodiment, the electronic data files may be produced by the data source devices 109 at a predetermined event or time (e.g. an end of a business day). Alternatively, the electronic data files may be produced on an hourly, weekly, or at any other appropriate time interval.

[0028] One or more data file formats may be associated with each of the data source devices 109. Each of the produced electronic data files may be associated with a unique data file identifier. Alternatively, each group of data files produced by a single data source device 109 (e.g., data source device 109-1) may be associated with a unique data source identifier associated with that data source device (e.g., data source device 109-1). One or more of the data source devices 109 may be uniquely configured to produce the one or more electronic data files in accordance with data subscription unit 101 of the data conversion and distribution system 100.

[0029] An example data subscription unit 101 of the data conversion and distribution system 100 of FIG. 1A is depicted in FIG. 2. The data subscription unit 101 may include at least one data interface 201 communicatively coupled via network 110 to plurality of data source devices 109. The data subscription unit 101 may be configured to obtain data having a plurality of data formats via the electronic data files produced by the one or more data source devices 109. The data subscription unit 101 may include one or more processors 209 (also referred to herein as processing component 209), logic 210 and a non-transitory memory 205 including instructions

206 and space to store subscription preferences. The subscription preferences may define the parameters of the communicative coupling between the data subscription unit 101 and the plurality of data source devices 109. In other words, the subscription preferences may define which data source devices 109 to connect to and communicate with, the type, volume and/or frequency with which data is pulled or received from said data source devices 109, and/or any other parameters related to the flow of data and information. The data subscription unit 101 may also include a data transmitter 207 configured to transmit the obtained data (having the plurality of data formats) via secure communication over network 108. Transmissions from the data transmitter 207 may be received by the virtual machine 103 of the data conversion and distribution system 100.

[0030] The data subscription unit 101 may, for example, via processor 209, receive subscription preferences, store the received subscription preferences in the non-transitory memory 205, and communicatively couple via the at least one data interface 201 of the data subscription unit 101 to one or more of the data source devices 109. In one embodiment, communicatively coupling via the at least one data interface 201 of the data subscription unit 101 to the data source devices 109 further includes sending a request (from the data subscription unit 101) to the data source devices 109 to receive data files related to a particular input or data, over a particular communication link, at a specified frequency. The data subscription unit 101 may then connect to the data source devices 109 by establishing a communication link between the data interface(s) 201 of the data subscription unit 101 and the data source device(s) 109 in network 110. The network 110 may be unsecured or secured and wired and/or wireless.

[0031] The data subscription unit 101 is said to be subscribed to a data source device 109 if a request transmitted to at least one data source device (e.g., data source device 109-1) among data source devices 109 is accepted and data and information is transmitted in accordance with the request from the data source device(s) 109 to the data subscription unit 101 via the network 110. In one embodiment, a request may specify the type and/or volume of data and information requested, the frequency at which it should be transmitted, as well as the communication protocol that should be used to transmit the data and information. For example, a request may requesting that one or more data source devices 109 transmits electronic data files regarding all sales activity relating to instrument or product X at the end of every business day in accordance

with a file transfer protocol (FTP) or secure file transfer protocol (SFTP). Alternative secure communication links may also be utilized.

[0032] In accordance with the received request, the respective data source device(s) 109 may generate one or more electronic data files containing only the requested information and transmit the requested data files at the specified frequency. The generated electronic data file(s) may then be transmitted to the data subscription unit 101 via data interface 201. In this manner, an embodiment of the data conversion and distribution system 100 may dictate receiving only the type and volume of data and information that is pertinent to supplementing and/or generating statistical information (e.g., data projections and sensitivities) related to one or more electronic data classes for which directly-related or historical information is sparse or unavailable. In this manner, the processing and memory requirements of the data conversion and distribution system 100 are maximized (i.e., by avoiding receiving irrelevant or voluminous data beyond what is needed or desired), particularly in embodiments where it is envisioned that millions of data requests and/or data files are received per day.

[0033] The electronic data files received by the at least one data interface 201 of the data subscription unit 101 may be in a variety of formats. For example, the data file formats may correspond to the specifications of each of the data source devices 109 from which the data files are received. Additionally, the data file formats may have different data transfer parameters, compression schemes, and the like. Furthermore, in some examples, the data file content may correspond to different forms of data, such as different currencies, date formats, time periods, and the like. In one embodiment, the data interface(s) 201 may receive a separate electronic data file for each request for information. In another embodiment, the data interface 201 may receive a single data file, corresponding to one or more requests for information, from each of the plurality of data source devices 109 to which it subscribes.

[0034] Thus, the frequency and volume of data which is provided to the data subscription unit 101 and the setup for a communication link may be arranged in accordance with the subscription preferences stored on the data subscription unit 101. The subscription preferences may be provided by a user device connected to the data conversion and distribution system 100 (either via a direct and/or remote connection to data subscription unit 101, or by way of any other input means of the data conversion and distribution system 100) and/or by an operator of the data

conversion and distribution system 100 itself. The preferences may be stored on the non-transitory memory 205 of the data subscription unit 101. Optionally, the data received via the data interface 201 may also be stored in the non-transitory memory 205 of the data subscription unit 101. In one embodiment, newly received data from the one or more data source devices 109 may be used to update, add to, or remove data already stored in the non-transitory memory 205 of the data subscription unit 101.

[0035] In one embodiment, the subscription preferences may be received by a data subscription preference receiver 203 specially configured to receive subscription preferences, and store and/or update subscription preferences in at least a portion of the non-transitory memory component 205 of the data subscription unit 101.

[0036] In one embodiment, after the data source devices 109 are subscribed to by the data subscription unit 101, the data may be automatically transmitted from the data source devices 109 to the data subscription unit 101 as the electronic data files are generated on the data source devices 109. In one embodiment, a predetermined event or time (e.g., the close of a business day or a predetermined time of day) may cause the data source device 109 to generate the data files for the data subscription unit 101.

[0037] In one embodiment, the data subscription unit 101 may further include one or more security protocols. The security protocols may include, for example, verification of one or more of the unique identifiers associated with the received electronic data files, including, for example the unique data file identifier and/or a unique data source identifier. For example, in one embodiment, the unique data source identifier may be utilized by the data subscription unit 101 to verify that it is receiving data files and information from the appropriate data source device 109. Such a system may be advantageous in preventing denial of service attacks and other malicious actions which are intended to harm the data conversion and distribution system 100 or the remote user device(s) 107 (e.g., by way of the data conversion and distribution system 100).

[0038] The data subscription unit 101 further includes a data transmitter 207 configured to transmit the data having the plurality of data formats via secure communication over a network 108. In one embodiment, a FTP or SFTP connection may deliver the received data files including the plurality of data formats to a virtual machine 103 of the data conversion and distribution system 100 via the data transmitter 207.

[0039] As illustrated in FIG. 3, an example virtual machine 103 of the system of FIG. 1A may include non-transitory memory 303 storing machine readable instructions 304, and one or more processors 305 (also referred to herein as processing component 305) including processor logic 306. The virtual machine 103 is communicatively coupled to the data subscription unit 101. The virtual machine 103 may also include a data receiver module 307, a data unification module 309, a data conversion module 311, and/or a data transmission module 315. Although the virtual machine 103 is illustrated in FIG. 1A as a single machine (e.g., a server), in some examples, the virtual machine 103 may include one or more servers.

[0040] The data receiver module 307 may be configured to receive electronic data having the plurality of data formats from the data subscription unit 101 via an optionally secure communication over the network 108. Once the data receiver module 307 receives the data having the plurality of data formats, it may transfer the data from the data receiver module 307 to the data unification module 309 for processing.

[0041] The data unification module 309 may be configured to receive data having the plurality of data formats from the data receiver module 307. Upon receiving the data having the plurality of data formats, the data unification module 309 may at least one of reformat, aggregate, decompress, cleanse and/or unpack the data having the plurality of data formats in order to generate unified data. Reformatting the data having the plurality of data formats may include analyzing the received data to identify its data type, and converting the received data into data having a predefined data format or type. For example, reformatting may involve converting data having different formats (e.g., comma separated variables (CSV), extensible markup language (XML), text) into data having a single format (e.g., CSV).

[0042] In one embodiment, the data having a plurality of data formats (and originating from a plurality of data source devices 109) may be aggregated. Aggregation may involve combining data and/or a plurality of electronic data files from one or more data sources into a single compilation of electronic data (e.g., one electronic data file) based on certain parameters and/or criteria. For example, in one embodiment, data may relate to a particular product or instrument, and recent observations including information regarding transaction counts, quote counts, transaction volume or price histories from a variety of dates and/or time periods may be combined or aggregated for each particular product or instrument.

[0043] At least a portion of the data having the plurality of data formats may be received by the data unification module 309 in a compressed format (which means that the data has been encoded using fewer bits than was used in its original representation). The data received in compressed format may be decompressed by the data unification module 309, which involves returning the data to its original representation for use within the virtual machine 103. For example, “zipped” data files (which refer to data files that have been compressed) may be “unzipped” (or decompressed) by the data unification module 309 into electronic data files having the same bit encoding as they did prior to their being “zipped” (or compressed).

[0044] Cleansing the data may include scanning and/or analyzing a volume of raw data and identifying and removing any data and information deemed incorrect, out-of-date, redundant, corrupt, incomplete and/or otherwise not suitable or non-useful for purposes of supplementing the sparse data set and/or performing statistical analyses for the sparse data set. It is envisioned that the volume of raw data may include data and information pertaining to millions (even tens of millions) of products or instruments. Thus, performing the cleansing function will substantially reduce the volume of data and information that is subject to subsequent functions described herein (e.g., aggregating, unpacking, reformatting, decompressing, etc.). As a result, fewer system resources will be required to perform any of these subsequent functions. In this manner, the cleansing function operates to improve overall system operating efficiency and speed.

[0045] Removing data that is determined to be unsuitable or non-useful from the raw data may involve a filtering function that separates the suitable and useful data from the unsuitable and non-useful data, and then forwards only the suitable and useful data for further processing. The data deemed unsuitable or non-useful may be deleted, stored in a dedicated storage location and/or otherwise disposed of. Cleansing the data may also include aligning data received from multiple sources and/or at multiple times, where aligning may involve assembling the data in a form that is suitable for processing by the data conversion module 311 (e.g., sorted according to a time sequence, grouped by category, etc.). In one embodiment, cleansing the data may also include converting data in one form (as opposed to type or format) into data having a standardized form that is usable by the data conversion module 311 (e.g., currency conversion).

[0046] Unpacking the data may or may not include one or more of the decompressing, cleansing, aggregating, and/or other functions described above. Alternatively or additionally,

unpacking may involve opening one or more data files, extracting data from the one or more data files, and assembling the extracted data in a form and/or format that is suitable for further processing. The sequences for opening and/or assembling the data may be predefined (for example, data may be opened/assembled in a sequence corresponding to timestamps associated with the data).

[0047] One or more of the functions discussed above (including, for example, reformatting, aggregating, decompressing, cleansing, and unpacking) as being carried out by the data unification module 309 may be performed in any suitable order or sequence. Further, one or more of these functions may be performed in parallel, on all or on portions of the received data. Still further, one or more of these functions may be performed multiple times. Collectively, one or more of these functions may be performed by the data unification module 309 (on the received data having a plurality of data formats) to ultimately generate the unified data (e.g., data having similar data characteristics (e.g., format, compression, alignment, currency, etc.)). The data unification module 309 may also perform additional and/or alternative functions to form the unified data.

[0048] Since the data unification module 309 may be separate and upstream from remote user devices 107, the processing functions discussed above are performed external to the remote user devices 107. Accordingly, the remote user devices 107 are able to receive electronic data from multiple data sources 109 in a unified form (and/or unified format) without having performed such aggregating and reformatting functions. Additionally, the data source devices 109 no longer have to reformat the data it generates prior to transmitting it to the data conversion and distribution system 100, as the data subscription unit 101 and the virtual machine 103 are able to receive and process data having any of the plurality of data formats.

[0049] At least a portion of the unified data may be stored in the memory 303 of the virtual machine 103. The memory 303 of the virtual machine 103 may be modular in that additional memory capabilities may be added at a later point in time. In one embodiment, it is envisioned that a virtual machine 103 of a data conversion and distribution system 100 may be initially configured with approximately 15 GB of disk space and configured to grow at a rate of 1.5 GB per month, as the virtual machine 103 receives and then stores more data from the data

subscription unit 101, although any initial amount of disk space and any growth rate may be implemented.

[0050] The solutions described herein utilize the power, speed and precision of a special purpose computer system configured precisely to execute the complex and computer-centric functions described herein. As a result, a mere generic computer will not suffice to carry out the features and functions described herein. Further, it is noted that the systems and methods described herein solve computer-centric problems specifically arising in the realm of computer networks so as to provide an improvement in the functioning of a computer, computer system and/or computer network. For example, a system according to the present disclosure includes an ordered combination of specialized computer components (e.g., data subscription unit, virtual machine, etc.) for receiving large volumes of data having varying data formats and originating from various data sources, reformatting and aggregating the data to have a unified format according to preferences, and then transmitting the unified data to remote user devices. As a result, the remote user devices only receive the type and volume of information desired and the remote user devices are freed from performing the cumbersome data processing and conversion functions accomplished by the specialized computer components.

[0051] The unified data (provided by data unification module 307) may be accessed by or transferred to the data conversion module 311. The data conversion module 311 is configured to execute one or more statistical processes (e.g., statistical modeling, algorithms, etc.) using the unified data to generate at least one of data sensitivities, projected data, and/or any other statistical analyses information based on the unified data. In one embodiment, the data conversion module 311 may be configured to model and produce projected data based on the unified data, and data sensitivity information may be determined based on the projected data. In this manner, the data conversion module 311 is able to produce projected data and data sensitivities (and other statistical analyses information) for data classes without sufficient direct data to generate said projections, sensitivities, etc. (e.g., data classes having sparse electronic data). It may also be appreciated that data projections and data sensitivities may be reviewed according to archived data, to adjust modeling used by the statistical algorithm(s).

[0052] One example of a sparse electronic data set includes electronic transactional data associated with liquidity indicators. Participants in such an industry (including portfolio

managers, analysts, regulatory compliance teams, etc.) may seek information related to whether a product or instrument has sufficient liquidity. Existing computer systems offer variations of “liquidity scoring” which largely depends on a counted number of data points (i.e., dealer sources) that have been observed. However, in illiquid markets, directly observable data points relating to transactional and quote information may be scarce. For example, in some fixed income markets, less than 2% of the issued instruments are a part of a transaction on a given day. As a result, directly observable data points relating to transaction and quote information is sparse, thereby forming a sparse electronic data set.

[0053] Accordingly, a data conversion and distribution system according to the current disclosure provides a solution for these types of data classes having sparse electronic data sets. As described above, the solution comes in the form of specially configured computer components, including a data subscription unit and a virtual machine, that collectively, receive any amount of data according to preferences, the data having varying data formats and originating from a variety of data sources, reformat and aggregate the data, and generate unified data files that may be run through statistical algorithms to generate statistical data and information for the sparse data classes.

[0054] Some portions of the description herein describe the embodiments in terms of algorithms and symbolic representations of operations on information. These algorithmic descriptions and representations are commonly used by those skilled in the data processing arts to convey the substance of their work effectively to others skilled in the art. These operations, while described functionally, computationally, or logically, are understood to be implemented by computer programs or equivalent electrical circuits, microcode, or the like. Furthermore, it has also proven convenient at times, to refer to these arrangements of operations as modules, without loss of generality. The described operations and their associated modules may be embodied in specialized software, firmware, specially-configured hardware or any combinations thereof.

[0055] Additionally, certain embodiments described herein may be implemented as logic or a number of modules, components, or mechanisms. A module, logic, engine, component, or mechanism (collectively referred to as a “module”) may be a tangible unit capable of performing certain operations and is configured or arranged in a certain manner. In certain exemplary embodiments, one or more computer systems (e.g., a standalone, client, or server computer

system) or one or more components of a computer system (e.g., a processor or a group of processors) may be configured by software (e.g., an application or application portion) or firmware (note that software and firmware may generally be used interchangeably herein as is known by a skilled artisan) as a module that operates to perform certain operations described herein.

[0056] In various embodiments, a module may be implemented mechanically or electronically. For example, a module may include dedicated circuitry or logic that is permanently configured (e.g., within a special-purpose processor) to perform certain operations. A module may also include programmable logic or circuitry (e.g., as encompassed within a specially-purposed processor or other programmable processor) that is configured (e.g., temporarily) by software or firmware to perform certain operations.

[0057] Accordingly, the term module should be understood to encompass a tangible entity, be that an entity that is physically constructed, permanently configured (e.g., hardwired), or temporarily configured (e.g., programmed) to operate in a certain manner and/or to perform certain operations described herein. Considering embodiments in which modules or components are temporarily configured (e.g., programmed), each of the modules or components need not be configured or instantiated at any one instance in time. For example, where the modules or components include a specially purposed processor configured using software, the specially purposed processor may be configured as respective different modules at different times. Software may accordingly configure the processor to constitute a particular module at one instance of time and to constitute a different module at a different instance of time.

[0058] FIG. 4 is a flowchart of one example statistical algorithm that may be used in connection with the data conversion module 311 of FIG. 3 and is related to providing liquidity indicator statistics. Liquidity may be defined as the ability to exit a position at or near the current value of a product or instrument. For purposes of this disclosure, a product or instrument shall refer to any asset, whether tangible or electronic, that may be purchased, sold, offered, exchanged or otherwise made the subject of a transaction). In some embodiments, a product or instrument may refer to a consumer good, while in others, it may refer to a securities or similar assets.

[0059] The data conversion and distribution system 100 described herein may be used, in one exemplary and non-limiting embodiment, to generate liquidity indicator statistics for fixed income instruments which, as discussed above, may not be the object of active transactional activities. Fixed income instruments may include individual bonds, bond funds, exchange traded funds (ETFs), certificates of deposits (CDs), money market funds and the like. This approach to measuring liquidity, however, is not limited to fixed income securities, and is applicable to other types of instruments, including but not limited to, equities, options, futures, and other exchange-listed or OTC derivatives. Illiquid markets such as fixed income markets have limited transactional activity. For example, less than 2% of the outstanding instruments in fixed income markets may be the subject of transactional activity on any given day. Thus, data such as market depth is insufficient to construct an accurate assessment of an instruments' statistical liquidity. Accordingly, in one embodiment, a statistical algorithm of FIG. 4 may be used to estimate statistical indicators of an instrument's liquidity (e.g., "liquidity indicators") based on the influence of features on the ability to exit a position at or near the current value of the instrument. The statistical algorithm of FIG. 4 may be run on a specialized liquidity engine of the data conversion module 311. The liquidity engine may be configured specifically for providing statistical liquidity indicators.

[0060] In the statistical algorithm of data conversion module 311 shown in FIG. 4, features of the buyers, sellers, and asset may be used to determine the ability to electronically transact a particular instrument. Features may include asset class, sector, issuer, rating (investment grade, or high-yield), maturity date, amount outstanding, issue date, and index constituent, number of quotes, number of transactions, number of holders, number of buyers and sellers, transaction volume, tighter bid/ask spreads, liquidity premiums and the like. The influence of features on the transaction volume may be determined by applying a statistical algorithm comparing historical data regarding the features to historical information regarding the transaction volume. The results of the statistical algorithm may be applied to information about the current features of the instrument in order to project the future transaction volume, liquidity and the like.

[0061] The statistical algorithm of FIG. 4 may include a number of pre-modeling steps 415, including receiving unified data 401 that may include data quote counts, transaction counts, and transaction volumes values corresponding to a time window. The statistical algorithm may then

determine timing information 403. In particular, the received time window may be broken into time periods. For example, the time window may include 84 business days and may be subdivided into 4 time periods of 21 days each.

[0062] The data and information in each of the time periods may be used to derive price volatilities 405 for each instrument. To derive the price volatilities, a time horizon may be defined. In one embodiment, the time horizon may depend on the time to maturity. For example, if the days to maturity is greater than 53, then the time horizon may be set to 63 days, and if the days to maturity is less than or equal to 53 days, then the time horizon may be set to the days to maturity plus 10 days. Once the time horizon is defined, the price volatility 405 may be derived by comparing the bid price for each instrument in the time horizon in sequential order from the most recent bid to the earliest bid in the time horizon. In one embodiment, the comparison may include calculating the average absolute log price change for each sequential pair of bids. Determination of the price volatilities may include use of stored unified data or unified data that includes historical trade information.

[0063] The statistical algorithm of FIG. 4 may also calculate holders data for each asset class 407. For example, the statistical algorithm may calculate the median holders over two time periods (e.g., each time period spanning 42 production days).

[0064] The statistical algorithm of FIG. 4 may include additional filtering steps 409 for identifying instruments which are eligible to receive a liquidity score. In this example, instruments may refer to securities or any other similar product. The statistical algorithm may further include a filtering rule set which is applied to instruments. For example, the filtering rule set may specify that a particular instrument be “ignored”. A liquidity score may not be calculated for an “ignored” instrument. The filtering rule set may also specify that an instrument that is actively evaluated and released by the organization implementing the data conversion and distribution system be ignored.

[0065] The statistical algorithm of FIG. 4 may determine a list of inputs 411 for use in modeling. These inputs may include one or more of an instrument identifier, issue date, quote count, trade count, trade volume, amount outstanding, issuer identifier, financial Boolean, investment grade Boolean, and the like. These inputs may be obtained from the unified data provided by data unification module 309.

[0066] Prior to calculating the liquidity indicators, the algorithm may bucket and sort a number of instruments 413 according to the price volatilities of each instrument. The instruments may be bucketed in accordance with their different durations. Within each bucket, the instruments may be sorted based on their volatility value. For example, the system may create 40 distinct buckets for each list of instruments, where the instruments are bucketed by their durations. Within each bucket, the instruments may be sorted by their price volatilities. In one embodiment, near-zero or zero-valued price volatilities may be replaced with the minimum non-zero volatility. Similarly, if an entire bucket having non-zero valued volatilities is included, a predetermined percentage (e.g., the lowest ten percent (10%)) of the volatilities may be replaced with the first volatility value found after the predetermined percentage (e.g., the lowest ten percent (10%)).

[0067] The statistical algorithm of FIG. 4 may include modeling steps 433 involving one or more non-regression models 425 and one or more regression models 417. The one or more models 417, 425 of modeling step 433 may be run for each type of instrument independently. For example, the one or more regression models 417 may be run on investment grade bonds (which have a low risk of default) independently from running the one or more regression models on high-yield bonds (which have lower credit ratings and a higher risk of default).

[0068] In one embodiment, at least one of the one or more regression models 417 is a linear multifactor regression model. The one or more regression models 417 may be utilized to generate correlation sensitivities (data sensitivities) between factors or attributes (an X-side of the regression) and the transaction volume (a Y-side of the regression) of an instrument 421. The correlation sensitivities (data sensitivities) may then be used to project future trade volumes 423.

[0069] In one embodiment, two regression models, Models A and B, may be utilized to generate correlation sensitivities (data sensitivities) or beta-values, between factors (attributes) and transaction volume. Model A may use one or more factors (attributes) related to the transaction volume, quote count, transaction count, amount outstanding (AMTO), years since issuance (YSI), financial Boolean, holders data (calculated above in step 407), bond price and the like for the X-side of the regression 419. Model B may use factors (attributes) related to the issuer transaction volume, issuer quote count and transaction count, AMTO, financial Boolean,

holders data (calculated above in step 407), bond price and the like for the X-side of the regression 419. The years since issuance may be calculated as the difference in the number of days between the issue date and the current production date and dividing the difference by 365. Both Model A and Model B may use the most recent time period (calculated above in step 403) for the Y-side of the regression 419. In one embodiment, the X-side factors (attributes) for the transaction volume variable may be weighted so that the transaction volume values of the data set sums to the total transaction volume. Data and information related to these factors (attributes) may be obtained by the pre-modeling processing steps 415 described above.

[0070] The regression models 417 may generate correlation sensitivities or beta-values for the factors 421. For example, the two regression models, Models A and B, may be performed using the X-side and Y-side factors described above. The resulting correlation sensitivities 421 (i.e., data sensitivities) or beta-values may be indicative of the correlation between the X-side factors and the Y-side trading volume. In particular, the generated beta-values may indicate the correlation between the transaction volume, quote count and trade count, amount outstanding, years since issuance, financial Boolean, investment grade Boolean, holders, transformed bond price variable (e.g., may be defined by equation: $(\text{bond price} - 100)^2$), and the trading volume. In one embodiment, four separate sets of beta-values may be generated, as models A and B may be run separately for investment grade and high-yield bonds, as they are sensitive to different factors.

[0071] The correlation sensitivities or beta-values may then be used along with data and information corresponding to the factors in a new data set of the model to generate a projected volume 423. The new data set may be a portion of the unified data.

[0072] In one embodiment, alternative statistical models which do not use regression (non-regression models 425) may be used in combination with the regression models 417. In one embodiment, a model 425 with no regression step may calculate the projected volume as a weighted sum average of the transaction volume from a set number of time periods 427. In another embodiment, a model 425 with no regression step may calculate the projected volume as the maximum of average accumulative volume of all of the previous days up to the current day in a time period 427. In yet another embodiment, a model 425 with no regression step may calculate the projected volume as the average volume across a time period 427.

[0073] In certain embodiments, a seasonal adjustment may be applied to the projected volume from the regression or non-regression models (425, 417) of projected volume. Additionally, one or more algorithms may be run on the projected volumes to remove the effects of regression linkage.

[0074] Various post-modeling steps 439 may be taken by the statistical algorithm of data conversion module 311. The outputs from the one or more regression and non-regression models (425, 417) applied on the unified data may be utilized to determine a projected volume and a projected dollar volume for any bond 429. In one embodiment, the projected volume is the maximum volume from all applicable models. The projected dollar volume may be calculated as the projected volume * BidPrice/100. The BidPrice may be indicative of the price a buyer is willing to pay for the instrument. The projected dollar volume may be subject to a minimum dollar volume rule such that if the projected volume is less than 1000 and the amount outstanding is less than 1000 but not equal to zero, the projected dollar volume may be set to the AMTO*BidPrice/100. Alternatively, if the projected volume is less than 1000 and the amount outstanding is greater than 1000, the projected dollar volume is set to 1000*BidPrice/100.

[0075] After a projected dollar volume is generated for each instrument (step 429), the algorithm may generate an Amihud ratio value 431. The Amihud ratio is indicative of illiquidity and is commonly defined as a ratio of absolute stock return to its dollar volume averaged over a time period. The Amihud ratio value may be calculated by identifying the volatility of each instrument (see step 405), and dividing the volatility by the max projected dollar volume across all the models (see step 429).

[0076] The models 425, 417 (collectively, 433) may output a number of measures that are available for use by downstream products. These outputs may include the active trading estimate (the maximum dollar volume of the non-regression models), the potential dollar volume (maximum dollar volume of the regression models), the Projected Trade Volume Capacity (the maximum dollar volume across all of the regression and non-regression models), the volatility, and the Amihud ratio value.

[0077] The outputs from the models 433 may also be used to assign scores that allow for the comparison of instruments. Those instruments having a low Amihud ratio value may be given a high score indicating they are the more liquid instrument. Those instruments having a high

Amihud ratio value may be given a low score indicating they are a less liquid instrument. Scores may be determined based on an instrument's percentile rank in comparison with the universe size (the number of unique Amihud ratio values). The instruments in each category may be ranked in a list. In one example, the list may be separated into ten sections, where the first 10% having the highest Amihud scores are assigned a score of 1, the second 10% having the next highest Amihud scores are assigned a score of 2, and so forth.

[0078] The statistical algorithm may also determine the liquidity ratio 435, which is a liquidity indicator (described further below). The liquidity ratio 435 is an estimate of the market price response per dollar transacted in an instrument. The liquidity ratio 435 may be defined as the projected future potential price volatility divided by the projected future potential transaction volume (determined in step 429). The liquidity ratio may be a normalized value (as each instrument is normalized by its projected future potential transacting volume), and thus allows for the direct comparison of instruments within a given category 437.

[0079] The statistical algorithm may determine a liquidity score per category 437. Categories for ranking the instruments may include one or more of all bonds, same asset class, same sector, same issuer, similar duration in asset class, similar yield to maturity in asset class, and similar amount outstanding bonds in asset class. The all bonds category may include every instrument that received an Amihud value for the given production date, across all asset types (corporate, municipal, structured, agency, etc.).

[0080] The same asset class category may cover instruments having the same asset class. In other words, corporate instruments may be compared to corporate instruments and municipal bond instruments may be compared to municipal bond instruments. The same sector category may cover instruments categorized with the same market sector. The same issuer category may cover instruments assigned to the same issuer id. The same duration in asset class category may cover instruments with similar duration ranges within the same asset class. The duration ranges may be derived by sorting the instruments by their duration value, breaking the sorted list into ten equally weighted ranges, and assigning each of the ten equally weighted ranges a score. The similar yield to maturity in asset class category may cover instruments with similar yield to maturity ranges within the same asset class. The yield to maturity ranges may be derived by sorting the instruments by their yield to maturity value, breaking the sorted list into ten equally

weighted ranges, and assigning each of the ten equally weighted ranges a score. The similar outstanding bonds in asset class category may cover instruments with similar amount outstanding ranges within the same asset class. The amount outstanding ranges may be derived by identifying unique amount outstanding values per asset class, sorting the instruments by their amount outstanding values per asset class, breaking the sorted list into ten equally weighted ranges, and assigning each of the ten equally weighted ranges a score.

[0081] The output from these models (active trading estimate, the potential dollar volume, the Projected Trade Volume Capacity, the Projected Volatility, the Amihud ratio value, and the liquidity scores) are examples of liquidity indicators. Scoring, categorical information, outputs from the models, liquidity indicators, may be stored on the memory component 303 of the virtual machine 103, the data distribution device 105, and made available for downstream products and applications on a remote user device 107.

[0082] The output from the data conversion module 311 (including, for example, regression and non-regression models (425, 417), liquidity indicators, scoring, categorical information and the like) may be transmitted via the data transmission module 315 of the virtual machine 103 to the data distribution device 105 via one or more secure communications over network 108.

[0083] An example data distribution device 105 of the system of FIG. 1A is depicted in FIG. 5. The data distribution device 105 may include one or more processors 503 (also referred to herein as processing component 503) including processor logic 504. The data distribution device 105 may include at least one data distribution receiver 505 configured to receive information from the virtual machine 103. The data distribution device 105 may include non-transitory memory 501 including instructions 502 to store the outputs from the regression and non-regression models (425, 417), liquidity indicators, scoring, categorical information, and/or any other derived statistical data or information from the virtual machine 103.

[0084] The data distribution device 105 may include at least one data distribution interface 507 configured to provide secure communications with at least one remote user device via network 106. The non-transitory memory 501 of the data distribution device 105 may also be configured to store predefined settings for one or more remote user devices 107. The data distribution device 105 may be further configured to receive a request from one or more remote user devices 107 at data distribution receiver 505. The request may detail which portion of the

stored information on the data distribution device 105 the respective remote user device 107 indicates to receive. The data distribution device 105 may send the requested portion of the stored information to the remote user device 107 responsive to receiving the request. For example, a remote user device 107 may request that the data distribution device 105 only transmit liquidity indicators for instrument X to the remote user device 107. Transmissions from the data distribution device 105 to the remote user devices 107 via the network 106 may involve FTP and a structured query language (SQL) loader, or any other suitable means. The contents of the request may form the predefined settings that are stored on the non-transitory memory 501 of the data distribution device 105.

[0085] An example remote user device is illustrated in FIG. 6. As illustrated in FIG. 6, remote user device 107 may include a non-transitory memory 601 storing machine readable instructions 602, one or more processors 603 (also referred to herein as processing component 603) including processor logic 604, a data distribution receiver interface 605, a user information interface 607, a market data source interface 609, and/or a user display interface 611. The data distribution receiver interface 605 may be specially configured to be communicatively coupled to the data distribution device 105 via network 106. For example, in one embodiment, the remote user device 107 may be specially configured to perform certain data processes, contain an up-to-date version of a web browser associated with system 100, and have an Internet connection capable of communication with system 100. The remote user device 107 may have an account with the service provider of the data conversion and distribution system 100. The remote user device 107, and, more specifically the data distribution receiver interface 605, may establish a secure connection with the data distribution device 105. The secure connection may be mediated by a password portal on a web-service, a secured application, biometrics device(s), and the like. Additional security measures which allow for encrypted communications (such as industry standard secured hypertext transfer protocol (HTTPS), secure socket layer (SSL) certificates, and the like) may also be used. Although a single remote user device 107 is discussed, a plurality of remote user devices 107 may be used with the data conversion and distribution system 100.

[0086] Each remote user device 107 may be configured to receive, via the data distribution receiver interface 605, at least one of the data sensitivities, projected values, and other information stored on the data distribution device 105. The remote user device 107 may also be

configured to receive user input data via the user information interface 607 and current market data via the market data source interface 609. The market data source interface 609 may be configured to receive market data from computer systems associated with exchanges, regulators and the like. In other embodiments, the market data source interface 609 may simply be a data source interface, configured to receive any type of form of data pertinent to any industry. The remote user device 107 may also be configured to generate supplementary projected data based on the received at least one of the data sensitivities and the projected data, the user input data and current market data. The projected data may include one or more of the projected volume, projected dollar volume, Amihud ratio, liquidity ratio and liquidity score per category. The supplementary projected data may include one or more of a projected market price impact and a projected days to liquidate.

[0087] Processing component 603 of each of the remote user devices 107 and processing component 503 of the data distribution device 105 may work in unison to generate supplemental projected data including a projected market price impact and a projected days to liquidate. For example, in one embodiment, a user of the remote user device 107 may upload and transmit data to the data distribution device 105. The uploaded and transmitted data may include the sparse data class and information relating thereto, such as product data, position data, instrument data, portfolio data, etc. The data distribution device 105 may receive and store the data from the remote user device 107. One or more algorithms stored on the memory component 501 of the data distribution device 105 may be executed to generate the supplemental projected data. Input to the one or more algorithms may include, for example, the data received from the remote user device 107, output from the data conversion module 311 (e.g., liquidity indicators, scoring, categorical information, and/or any other derived statistical data or information), data previously stored on the data distribution device 105, and/or other data and information relevant to the implementation. The supplemental projected data may then be transmitted from the data distribution device 105 to the remote user device 107. The remote user device 107 may receive and/or store the supplementary projected data from the data distribution device 105. The projected market price impact may be defined as the projected effect that a market participant will have when an instrument is bought or sold. It may be represented as a percentage. The projected days to liquidate may be defined as the projected days it would take to liquidate an

instrument given the position size of the instrument. In particular, a user of one of the remote user devices 107 may input a targeted market price impact via user information interface 607. The remote user device 107 may then retrieve projected data, data sensitivities, current market data, and other information related to the instrument. Using the obtained information the remote user device 107 (working with the data distribution device 105) may generate an estimate of the days to liquidate needed to achieve the targeted market price impact. Similarly, the remote user device 107 may receive from a user (via interface 606) a targeted projected days to liquidate. Using information obtained from the remote user device 107 and the data distribution device 105, the remote user device 107 and/or the data distribution device 105 may generate a measure of the projected market price impact given the targeted projected days to liquidate.

[0088] The supplemental projected data (including the projected market price impact and the projected days to liquidate) may take into account the impact of position size on liquidating an instrument. For example, two investors may hold the same instrument at varying positions: Investor A may have a \$1 million position and Investor B may have a \$100 million position. If the projected trading volume capacity is estimated to be \$10 million per day, it is reasonable to conclude that Investor A's position may be liquidated in one trading day, and Investor B's position may take longer to liquidate. Accordingly, the projected days to liquidate may take into account the projected trading volume capacity and position size. Additionally, there may be a time-dependent cost associated with exiting a position over the course of multiple days, as market conditions may change and influence the price of the asset. Thus, the projected market price impact may use the volatility estimates (used in the generation of the liquidity ratio), along with other variable considerations such as bid-ask spread and evaluated price of the security, to determine the impact on the market price based on how many days the investor uses to liquidate their position.

[0089] The remote user devices 107 may also display at least one of the projected data, supplementary projected data, user input data and current market data via the user display interface 611. The user display interface 611 may further include a graphical user interface (GUI), application programming interface (API) and the like. The remote user device 107 may be configured to receive user graphical user interface (GUI) preference data from a user of the system via interface 607. Using the received user GUI preference data, the remote user device

107 may extract information including at least a portion of the at least one of the projected data and the supplementary projected data, data sensitivities, and current market data from the memory 601 of the remote user device 107 and/or memory 501 of the data distribution device 105. The extracted information may then be displayed on the graphical user interface of the user display interface 611 in accordance with the user GUI preference data.

[0090] FIG. 7 illustrates an exemplary GUI 700 of the user display interface 611 of FIG. 6. In some examples, the GUI 700 may be present on a webpage accessed by the user of the remote user device 107. The GUI 700 may include a first section displaying instrument information 701 including, for example, the instrument title, a brief description, and the like.

[0091] The GUI 700 may also contain means for providing feedback to an operator of the data conversion and distribution system. Selection of the feedback icon 707 by the user may provide a pop-up window, link to a new tab or webpage, and the like which allows for communication with the system 100 for data conversion and distribution. Alternatively, hovering over the feedback icon 707 with a mouse, may display a phone number, email address, or chat service configured to aid in communication between the user of the remote user device 107 and the operator of the data conversion and distribution system 100.

[0092] A second section of the GUI 700 may include tabs 703 used to change the panels displayed in the GUI window. Tabs 703 may include transparency, best execution, liquidity, market data, evaluation history, instrument basics, puts/tender, call/sink/redemption, supplemental data, corporate actions, or any other desired tabs appropriate for the particular implementation. A selected tab may change color in order indicate to a user selection of the tab. Other panels displayed on the GUI window may be adjusted in accordance with the selected tab 703.

[0093] In the displayed embodiment, selection of the liquidity tab 703A displays at least five panels: a liquidity scores panel 709, a universe and liquidity rank panel 711, a score calculator panel 723, a comparable bonds panel 715, and a liquidity calculator panel 713. It is envisioned that additional or fewer panels may be visible upon selecting the liquidity tab 703A. The GUI 700 may also display information regarding the date at which data and information displayed in the GUI 700 was last updated 705.

[0094] The liquidity scores panel 709 may include information regarding the scores of each instrument when compared with the instruments in each categories, separated by category. Categories may include all bonds, same asset class, same sector, same issuer, similar duration bonds in an asset class, similar yield to maturity bonds in asset class, similar outstanding bonds in an asset class, etc. Each sub-panel 710 of the liquidity scores panel 709 may include the score 716, the category the score corresponds to 717, and an indicator 719. In one embodiment, selection of the indicator 719 may update the other panels and subpanels of the liquidity tab 703A. The selection of the indicator 719 may also display additional information related to the instrument and category chosen.

[0095] The universe and liquidity rank panel 711 may display information regarding the instrument's score in comparison with other instruments in the selected category 717. For example, the depicted example illustrates that a particular bond's score is more liquid than 18% (721) of the other bond scores within the same category 717 (asset class).

[0096] The score calculator panel 723 may display the projected data including the projected price volatility 725 and the projected volume capacity 727. The projected data may be depicted in numerical and/or graphical format 729, 731 for ease of use by the user. The score calculator panel 723 may also include the liquidity score 733, and a display of how the liquidity score may change over time 735 in graphical format.

[0097] The comparable bonds panel 715 may display a listing of instruments having the same issuer but with more favorable liquidity scores.

[0098] The liquidity calculator panel 713 may include an indication of whether a particular instrument is in a user's portfolio. The liquidity calculator may also include one or more fields 736 configured to receive user input. The fields 736 for user input may include position size, concentration, evaluated bid price, position market value, estimated transaction cost, stress level and/or any other information pertinent to the implementation. One or more of the fields may be updated automatically by the remote user computer device 107 based on either market data received from a market data source, or by other user input. Although textboxes configured for user input are depicted, alternate methods for receiving user input may be used, such as a scrollbar, selectable drop-down menu, and the like.

[0099] The liquidity calculator panel 713 may also include a display of the supplemental projected data including the projected days to liquidate 737 and the projected market price impact 739. It may also include a section depicting an estimation of the projected market price impact 743 given a number of target days to liquidate 741. Similarly, a section of the liquidity calculator panel 713 may also include an estimation of the projected days to liquidate 747 given a target market price impact 745.

[0100] Although exemplary sections and panels are depicted in FIG. 7, alternate configurations for the sections and panels are envisioned. For example, a graphical user interface may contain more or fewer sections and panels. Additionally, the sections and panels may be reorganized in any manner and display other pertinent information.

[0101] Additional panels 800 are depicted in FIG. 8. These additional panels 800 may be incorporated into the graphical user interface of FIG. 7. Alternatively, the additional panels 800 may be visible after selection of a separate tab 703 of the graphical user interface, or pop-up after selection of any element in FIG. 7. The additional panels 800 depicted in FIG. 8 include a liquidity coverage and distribution panel 801 which illustrates the total number of instruments 803 and a projected days to liquidate portfolio panel 805. The projected days to liquidate portfolio panel 805 may include user input fields 807 such as stress and targeted market price impact. After the user inputs the targeted market price impact by way of the sliding selector, the user input may be transmitted to the data distribution device 105. The data distribution device 105 and/or the remote user device 107 may work in unison to generate other projected values such as the projected days to liquidate. The projected days to liquidate may then be displayed in either the projected days to liquidate portfolio panel 805 in graphical or numerical form 809, or in the graphical user interface of FIG. 7 in the liquidity calculator panel 713 as element 747. Similar to the additional projected days to liquidate portfolio panel 805, it is envisioned that a graphical user interface may include a projected market price impact panel configured to receive from a user on a remote user device 107 the target days to liquidate. The user may input the target days to liquidate by way of a text-field, selection menu, selection boxes, slider or the like. The remote user device 107 may then transmit the target days to liquidate to the data distribution device 105 to obtain relevant data and information. The remote user device 107 and the data distribution device 105 may then work in unison to generate the projected market price impact.

[0102] Systems and methods of the present disclosure may include and/or may be implemented by one or more specialized computers including specialized hardware and/or software components. For purposes of this disclosure, a specialized computer may be a programmable machine capable of performing arithmetic and/or logical operations and specially programmed to perform the particular functions described herein. In some embodiments, computers may include processors, memories, data storage devices, and/or other specially-programmed components. These components may be connected physically or through network or wireless links. Computers may also include software which may direct the operations of the aforementioned components. Computers may be referred to with terms such as servers, personal computers (PCs), mobile devices, and other terms that may be interchangeable therewith, and any special purpose computer capable of performing the described functions may be used.

[0103] Computers may be linked to one another via one or more networks. A network may be any plurality of completely or partially interconnected computers, wherein some or all of the computers are able to communicate with one another. Connections between computers may be wired in some cases (e.g., via wired TCP connection or other wired connection) or may be wireless (e.g., via a Wi-Fi network connection). Any connection through which at least two computers may exchange data may be the basis of a network. Furthermore, separate networks may be able to be interconnected such that one or more computers within one network may communicate with one or more computers in another network. In such a case, the plurality of separate networks may optionally be considered to be a single network.

[0104] Each of the data source devices 109, data subscription unit 101, virtual machine 103, data distribution device 105, and remote user devices 107 may include one or more computing devices. The one or more computing devices may each include servers 301, processing components 209, 305, 503, 603 having logic 210, 306, 504, 604, memory components 303, 501, 601 having instructions 304, 502, 602, communications interfaces 315, 507, 607, 609, receivers 307, 505, 605, user displays 611 and/or the like.

[0105] Processing components 209, 305, 503, 603 may include, without being limited to, a microprocessor, a central processing unit, an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital signal processor (DSP) and/or a network processor. Processing components 209, 305, 503, 603 may be configured to execute processing

logic 210, 306, 504, 604 for performing the operations described herein. The processing components 209, 305, 503, 603 described herein may include any suitable special-purpose processing device or a processing device specially programmed with processing logic 210, 306, 504, 604 to perform the operations described herein.

[0106] Memory components 303, 501, 601 may include, for example, without being limited to, at least one of a read-only memory (ROM), a random access memory (RAM), a flash memory, a dynamic RAM (DRAM) and a static RAM (SRAM), storing computer-readable instructions 304, 502, 602 executable by processing components 209, 305, 503, 603. Memory components 303, 501, 601 may include any suitable non-transitory computer readable storage medium storing computer-readable instructions 304, 502, 602 executable by processing components 209, 305, 503, 603 for performing the operations described herein. Although one memory component 303, 501, 601 is illustrated in each of FIGS. 3, 5, and 6 in some examples, the one or more computer systems may include two or more memory devices (e.g., dynamic memory and static memory).

[0107] The one or more computing systems may include one or more communication interface interfaces 315, 507, 607, 609, and communication receivers 307, 505, 605, for direct communication with other computers and/or computer components (including wired and/or wireless communication) and/or for communication with network(s) 106, 108, 110 (FIG. 1A).

[0108] In some examples, the remote user devices 107 may include display devices (e.g., a liquid crystal display (LCD)). In some examples, computer system of a remote user device 107 may include one or more user interfaces 607, 611 (e.g., an alphanumeric input device, a touch sensitive display, a cursor control device, a loudspeaker, etc.).

[0109] In some examples, the one or more computer systems may include data storage devices storing instructions (e.g., software) for performing any one or more of the functions described herein. Data storage devices may include any suitable non-transitory computer-readable storage medium, including, without being limited to, solid-state memories, optical media and magnetic media.

[0110] The term "computer" shall refer to an electronic device or devices, including those specifically configured with capabilities to be utilized in connection with a data conversion and distribution system, such as a device capable of receiving, transmitting, processing and/or using

data and information in the particular manner and with the particular characteristics described herein. The computer may include a server, a processor, a microprocessor, a personal computer, such as a laptop, palm PC, desktop or workstation, a network server, a mainframe, an electronic wired or wireless device, such as for example, a telephone, a cellular telephone, a personal digital assistant, a smartphone, an interactive television, such as for example, a television adapted to be connected to the Internet or an electronic device adapted for use with a television, an electronic pager or any other computing and/or communication device specifically configured to perform one or more functions described herein.

[0111] The term “network” shall refer to any type of network or networks, including those capable of being utilized in connection with a data conversion and distribution system described herein, such as, for example, any public and/or private networks, including, for instance, the Internet, an intranet, or an extranet, any wired or wireless networks or combinations thereof.

[0112] The term “user interface” shall refer to any suitable type of device, connection, display and/or system through which information may be conveyed to and received from a user, such as, without limitation, a monitor, a computer, a graphical user interface, a terminal, a screen, a keyboard, a touchscreen, a biometric input device that may include a microphone and/or camera, a telephone, a personal digital assistant, a smartphone, or an interactive television.

[0113] The term “computer-readable storage medium” should be taken to include a single medium or multiple media that store one or more sets of instructions. The term “computer-readable storage medium” shall also be taken to include any medium that is capable of storing or encoding a set of instructions for execution by the machine and that causes the machine to perform any one or more of the methodologies of the present disclosure.

[0114] The term “or” may be construed in an inclusive or exclusive sense. Similarly, the term “for example” may be construed merely to mean an example of something or an exemplar and not necessarily a preferred means of accomplishing a goal.

[0115] While the present disclosure has been discussed in terms of certain embodiments, it should be appreciated that the present disclosure is not so limited. The embodiments are explained herein by way of example, and there are numerous modifications, variations and other embodiments that may be employed that would still be within the scope of the present invention.

CLAIMS

1. A data conversion and distribution system comprising:

a data subscription unit having at least one data interface communicatively coupled to a plurality of data source devices and configured to obtain data having a plurality of data formats from the plurality of data source devices and to transmit the data having the plurality of data formats via secure communication over at least one network; and

a virtual machine comprising one or more servers, a non-transitory memory, and one or more processors comprising machine readable instructions, the virtual machine communicatively coupled to the data subscription unit, the virtual machine further comprising:

a data receiver module configured to receive the data having the plurality of data formats from the data subscription unit via the secure communication over the at least one network;

a data unification module configured to reformat and aggregate the data from the data subscription unit to generate unified data responsive to receiving, at the receiver module, the data having the plurality of data formats, and further configured to at least one of decompress, cleanse, and unpack the data having the plurality of data formats;

a data conversion module configured to run the unified data through one or more statistical algorithms to generate at least one of data sensitivities and projected data based on the unified data; and

a data transmission module configured to transmit the at least one of the data sensitivities and the projected data to a data distribution device via one or more secure communications over the at least one network.

2. The system of claim 1, wherein the data distribution device further comprises:

a non-transitory memory configured to store the at least one of the data sensitivities and the projected data; and

at least one data distribution interface, each of the at least one data distribution interface configured to provide secure communications with at least one of one or more remote user devices.

3. The system of claim 2, wherein each of the one or more remote user devices further comprises:

- a non-transitory memory,
- one or more processors comprising machine readable instructions,
- a data distribution receiver interface communicatively coupled to the data distribution device;

- a user information interface;

- a market data source interface;

- a user display interface,

wherein each of the one or more remote user devices is further configured to:

- receive, via the data distribution receiver interface, at least one of the data sensitivities and the projected data from the data distribution device;

- receive, via the user information interface, user input data;

- receive, via the market data source interface, current market data;

- generate, via the one or more processors, supplementary projected data based on the received at least one of the data sensitivities and the projected data, the user input data and the current market data; and

- display, via the user display interface, at least a portion of at least one of the projected data and the supplementary projected data.

4. The system of claim 1, wherein the data subscription unit further comprises:

- a data preference receiver configured to receive subscription preferences; and

- a non-transitory memory configured to store the subscription preferences,

wherein the data subscription unit is further configured to communicatively couple the at least one data interface to at least one of the plurality of data source devices in accordance with the subscription preferences.

5. The system of claim 2, wherein each of the one or more remote user devices is configured to:

transmit, to the data distribution device, a request to receive a portion of the at least one of the data sensitivities and the projected data from the data distribution device; and

receive, from the data distribution device in a secure communication, the portion of the at least one of the data sensitivities and the projected data responsive to the request.

6. The system of claim 3, wherein each remote user device is further configured to:

receive user graphical user interface (GUI) preference data from a user of the device;

extract information comprising at least a portion of the supplementary projected data, the at least one of the data sensitivities and the projected data, and the current market data from at least one of the non-transitory memory of the remote user device and the non-transitory memory of the data distribution device based on the received user GUI preference data;

display the extracted information on a graphical user interface of the user display interface in accordance with the user GUI preference data.

7. The system of claim 1, wherein the projected data further comprises one or more of a projected volume, a projected dollar volume, an Amihud ratio, a liquidity ratio and a liquidity score per category.

8. The system of claim 3, wherein the supplementary projected data further comprises one or more of a projected market price impact and a projected days to liquidate.

9. A method for data conversion and distribution, the method comprising:

obtaining, by a data subscription unit via a secure communication over at least one network, data having a plurality of data formats from the data source devices;

transmitting, via the data subscription unit, the data having the plurality of data formats to a virtual machine comprising a data receiver module, a data unification module, a data conversion module, and a data transmission module;

receiving, by the data receiver module via secure communication over the at least one network, the data having the plurality of data formats from the data subscription unit;

reformatting and aggregating, via the data unification module, the data from the data subscription unit to generate unified data responsive to receiving, at the receiver module, the data having the plurality of data formats, said reformatting and aggregating comprising at least one of decompressing, cleansing, and unpacking the data having the plurality of data formats;

generating, via the data conversion module, at least one of data sensitivities and projected data, wherein said generating further comprises running the unified data through one or more statistical algorithms; and

transmitting, via the data transmission module, the at least one of the data sensitivities and the projected data to a data distribution device via one or more secure communications over the at least one network.

10. The method of claim 9, further comprising:

storing, by the data distribution device, at least one of the data sensitivities and the projected data; and

transmitting, by the data distribution device, at least one of the data sensitivities and the projected data to one or more remote user devices.

11. The method of claim 10, further comprising:

receiving, via a data distribution receiver interface of the one or more remote user devices, at least one of the data sensitivities and the projected data from the data distribution device, the one or more remote user devices comprising a non-transitory memory, one or more processors comprising machine readable instructions, an user information interface, a market data source interface, and a user display interface;

receiving, via the user information interface, user input data;

receiving, via the market data source interface, current market data;

generating, via the one or more processors, supplementary projected data based on the received at least one of the data sensitivities and the projected data, the user input data and the current market data; and

displaying, via the user display interface, at least a portion of the at least one of the projected data and the supplementary projected data.

12. The method of claim 9, further comprising:

- receiving, via the data subscription unit, subscription preferences;
- storing, at the data subscription unit, the received subscription preferences in a non-transitory memory; and
- communicatively coupling the at least one data interface of the data subscription unit to at least one of the plurality of data source devices in accordance with the subscription preferences.

13. The method of claim 10, further comprising:

- transmitting, from a remote user device from among the one or more remote user devices to the data distribution device, a request to receive a portion of the at least one of the data sensitivities and the projected data from the data distribution device; and
- receiving, at the remote user device from the data distribution device in a secure communication, the portion of the data sensitivities and the projected data responsive to the request.

14. The method of claim 11, further comprising:

- receiving, at each remote user device, user graphical user interface (GUI) preference data from a user of the device;
- extracting, at the respective remote user device, information comprising at least a portion of the supplementary projected data, the at least one of the data sensitivities and the projected data, and the current market data from at least one the respective remote user device and the data distribution device based on the received user GUI preference data;
- displaying, at the remote user device, the extracted information on a graphical user interface of the user display interface in accordance with the user GUI preference data.

15. The system of claim 9, wherein the projected data further comprises one or more of a projected volume, a projected dollar volume, an Amihud ratio, a liquidity ratio and a liquidity score per category.

16. The system of claim 11, wherein the supplementary projected data further comprises one or more of a projected market price impact and a projected days to liquidate.

17. A non-transitory computer-readable storage medium programmed to include instructions that, when executed by one or more processing devices, cause the one or more processing devices to perform the functions comprising:

- obtaining, by a data subscription unit via a secure communication over at least one network, data having a plurality of data formats from the data source devices;

- transmitting, via the data subscription unit, the data having the plurality of data formats to a virtual machine comprising a data receiver module, a data unification module, a data conversion module, and a data transmission module;

- receiving, by the data receiver module via secure communication over the at least one network, the data having the plurality of data formats from the data subscription unit;

- reformatting and aggregating, via the data unification module, the data from the data subscription unit to generate unified data responsive to receiving, at the receiver module, the data having the plurality of data formats;

- generating, via the data conversion module, at least one of data sensitivities and projected data, wherein said generating further comprises running the unified data through one or more statistical algorithms; and

- transmitting, via the data transmission module, the at least one of the data sensitivities and the projected data to a data distribution device via one or more secure communications over the at least one network.

18. The non-transitory computer-readable storage medium of claim 17, further comprising:

- storing, by the data distribution device, at least one of the data sensitivities and the projected data;

transmitting, by the data distribution device, at least one of the data sensitivities and the projected data to one or more remote user devices;

receiving, via a data distribution receiver interface of the one or more remote user devices, at least one of the data sensitivities and the projected data from the data distribution device, the one or more remote user devices comprising a non-transitory memory, one or more processors comprising machine readable instructions, a user information interface, a market data source interface, and a user display interface;

receiving, via the user information interface, user input data;

receiving, via the market data source interface, current market data;

generating, via the one or more processors, supplementary projected data based on the received at least one of the data sensitivities and the projected data, the user input data and the current market data; and

displaying, via the user display interface, at least a portion of at least one of the projected data and the supplementary projected data.