

SYSTEM AND METHOD FOR QUIZ GENERATION

TECHNICAL FIELD

The present disclosure relates to system and method for correlating documents, and more particularly relates to a system and method for quiz generation that enables grading.

BACKGROUND

Traditional approaches used of quiz generation require collection of essential information before framing questions for the quiz. The collection of essential information requires thorough study of multiple documents from which a set of questions are framed. To enhance the versatility and thus level of the questions in the quiz, the questions must be distinct from the previously generated or used set of questions. A substantial manual iteration is required to modify the language of the questions such that their context remains the same. This methodology of preparing question papers causes cringe-worthy moments for the teachers or other examining authorities.

In an attempt to reduce the manual intervention while preparing a question paper, various automated systems are developed. Further, various methodologies are developed for selection of essential information while framing questions. The conventional automated system lags in generating efficient questionnaire that covers different approaches to frame the questions. The conventional systems generate multiple questions which may be same or similar kind and thereby adding redundancy to the question paper. This obstructs in the levelling up the quality of question papers.

Therefore, there is a need to provide a system and method that is capable of solving the aforementioned problems of the conventional systems for generating questions and answers deriving context from a content.

SUMMARY

In view of the foregoing, a system and a method for quiz generation is disclosed. In some aspects, the quiz generation system includes a database and a processing circuitry. The database includes a first repository, configured to store a first set of documents and a second repository configured to store a second set of question-answer pairs. The processing circuitry is coupled to the database and is configured to generate a quiz.

For generation of the quiz, the processing circuitry is configured to generate a first text from the first set of documents. Further, the processing circuitry is configured to generate a first set of keywords

and the first set of question-answer pairs from the first text. Furthermore, the processing circuitry is configured to generate a second set of keywords from the second set of question-answer pairs. To generate the quiz, the processing circuitry is further configured to match the second set of keywords with the first set of keywords. Further, the processing circuitry is configured to select a number of question-answer pairs from the first set of question-answer pairs based on the matching of the second set of keywords with the first set of keywords.

In some aspects, to generate the first set of keywords, the processing circuitry is configured to pre-process the first text to generate a first string and process the first string to generate the first set of keywords. To generate the second set of keywords, the processing circuitry is configured to generate the second set of question-answer pairs to a second text, pre-process the second text to generate a second string and process the second string to generate the second set of keywords.

In some aspects, to match the second set of keywords with the first set of keywords, the processing circuitry is configured to encode each keyword from the first set of keywords to generate a first set of encoded vectors and encode each keyword from the second set of keywords to generate a second set of encoded vectors. Further, the processing circuitry is configured to determine the distance between each vector of the first set of encoded vectors and each vector of the second set of encoded vectors. Furthermore, the processing circuitry is configured to match the vectors from the first set of encoded vectors to the vectors of the second set of encoded vectors having distance less than a pre-defined threshold value.

In some aspects, the quiz generation system further includes a quiz memory that is coupled to the processing circuitry and configured to store the selected question-answer pairs from the first set of the question-answer pairs. The selected question-answer pairs are iteratively updated based on the matching of the second set of keywords with the first set of keywords. In some aspects, the quiz generation system is further configured to generate grades by evaluating a set of answers to a set of subjective questions.

BRIEF DESCRIPTION OF DRAWINGS

The drawing/s mentioned herein disclose exemplary aspects of the present disclosure. Other objects, features, and advantages of the aspect will be apparent from the following description when read with reference to the accompanying drawings. In the drawings, wherein like reference numerals denote corresponding parts throughout the several views:

The diagrams are for illustration only, which thus is not a limitation of the present disclosure, and wherein:

FIG. 1 represents a block diagram that illustrates a system configured to generate a quiz, according to an aspect of the present disclosure;

FIG. 2 represents a block diagram that illustrates a server of FIG. 1, according to an aspect of the present disclosure;

FIG. 3 represents a flow diagram that illustrates a method for quiz generation, according to an aspect of the present disclosure;

FIG. 4 represents a flow diagram that illustrates the steps in generation of the first set of question-answer pairs, according to an aspect of the present disclosure; and

FIG. 5 represents a flow diagram that illustrates a method for generating grades by the quiz generation system, according to an aspect of the present disclosure.

To facilitate understanding, like reference numerals have been used, where possible, to designate like elements common to the figures.

DETAILED DESCRIPTION OF THE DRAWINGS

The aspects herein and the various features and advantageous details thereof are explained more fully with reference to the non-limiting aspects that are illustrated in the accompanying drawings and detailed in the following description. Descriptions of well-known components and processing techniques are omitted so as to not unnecessarily obscure the aspects herein. The examples used herein are intended merely to facilitate an understanding of ways in which the aspects herein may be practiced and to further enable those of skill in the art to practice the aspects herein. Accordingly, the examples should not be construed as limiting the scope of the aspects herein. In the drawings, the sizes of components may be exaggerated for clarity.

Referring initially to the drawings, FIG. 1 represents a block diagram that illustrates a system 100 configured to generate a quiz (hereinafter interchangeably referred to as “system”), according to an aspect of the present disclosure.

In some aspects of the present disclosure, the system 100 may include a user device 102 and a server 104 communicatively coupled to each other via a communication network 106. The user device 102 may be configured to facilitate a user to input data, receive data, and/or transmit data within the system 100. Examples of the user device 102 may include, but are not limited to, a desktop, a notebook, a

laptop, a handheld computer, a touch sensitive device, a computing device, a smart-phone, and/or a smart watch. It will be apparent to a person of ordinary skill in the art that the user device 102 may include any device/apparatus that is capable of manipulation by the user.

In some aspects of the present disclosure, referring to FIG. 1, the user device 102 may include a user interface 110, a processing unit 112, a device memory 114, a quiz console 116 and a communication interface 118.

The user interface 110 may include an input interface for receiving inputs from the user. Examples of the input interface of the user interface 110 may include, but are not limited to, a touch interface, a mouse, a keyboard, a motion recognition unit, a gesture recognition unit, a voice recognition unit, or the like. Some aspects of the present disclosure are intended to include or otherwise cover any type of the input interface including known, related art, and/or later developed technologies. The user interface 110 may further include an output interface for displaying (or presenting) an output to the user. Examples of the output interface of the user interface 110 may include, but are not limited to, a digital display, an analog display, a touch screen display, a graphical user interface, a website, a webpage, a keyboard, a mouse, a light pen, an appearance of a desktop, and/or illuminated characters. The processing unit 112 may include suitable logic, instructions, circuitry, interfaces, and/or codes for executing various operations, such as the operations associated with the user device 102, or the like. In some aspects of the present disclosure, the processing unit 112 may utilize one or more processors such as Arduino or raspberry pi or the like. Further, the processing unit 112 may be configured to control one or more operations executed by the user device 102 in response to the input received at the user interface 110 from the user. Examples of the processing unit 112 may include, but are not limited to, an application-specific integrated circuit (ASIC) processor, a reduced instruction set computing (RISC) processor, a complex instruction set computing (CISC) processor, a field-programmable gate array (FPGA), a Programmable Logic Control unit (PLC), and the like. Some aspects of the present disclosure are intended to include or otherwise cover any type of the processing unit 110 including known, related art, and/or later developed processing units.

The user device 102 may further include the device memory 114, configured to store the logic, instructions, circuitry, interfaces, and/or codes of the processing unit 112, data associated with the user device 102, and data associated with the system 100. In some aspects, the device memory 114 may be configured to store metadata of data present at the server 104. Examples of the device memory 114 may include, but are not limited to, a Read-Only Memory (ROM), a Random-Access Memory (RAM), a flash memory, a removable storage drive, a hard disk drive (HDD), a solid-state memory,

a magnetic storage drive, a Programmable Read Only Memory (PROM), an Erasable PROM (EPROM), and/or an Electrically EPROM (EEPROM). Some aspects of the present disclosure are intended to include or otherwise cover any type of the device memory 114 including known, related art, and/or later developed memories.

In some aspects of the present disclosure, the user device 102 may further include the quiz console 116 configured as a computer-executable application, to be executed by the processing unit 112. The quiz console 116 may include suitable logic, instructions, and/or codes for executing various operations and may be controlled by the server 104. The one or more computer executable applications may be stored in the device memory 114. Examples of the one or more computer executable applications may include, but are not limited to, an audio application, a video application, a social media application, a navigation application, and the like.

In some aspects of the present disclosure, the user device 102 may further include a communication interface 118, configured to enable the user device 102 to communicate with the server 104 and other components of the system 100 over the communication network 106. Examples of the communication interface 118 may include, but are not limited to, a modem, a network interface such as an Ethernet card, a communication port, and/or a Personal Computer Memory Card International Association (PCMCIA) slot and card, an antenna, a radio frequency (RF) transceiver, one or more amplifiers, a tuner, one or more oscillators, a digital signal processor, a coder-decoder (CODEC) chipset, a subscriber identity module (SIM) card, and a local buffer circuit. It will be apparent to a person of ordinary skill in the art that the communication interface 118 may include any device and/or apparatus capable of providing wireless or wired communications between the user device 102 and the server 104.

The server 104 may be a network of computers, a software framework, or a combination thereof, that may provide a generalized approach to create the server implementation. Examples of the server 104 may include, but are not limited to, personal computers, laptops, mini-computers, mainframe computers, any non-transient and tangible machine that can execute a machine-readable code, cloud-based servers, distributed server networks, or a network of computer systems. The server 104 may be realized through various web-based technologies such as, but not limited to, a Java web-framework, a .NET framework, a personal home page (PHP) framework, or any web-application framework. The server 104 may be maintained by a storage facility management authority or a third-party entity that facilitates service enablement and resource allocation operations of the system 100. The server 104

may include the processing circuitry 120 and one or more memory units (hereinafter, collectively referred to and designated as “Database 122”).

The processing circuitry 120 may include suitable logic, instructions, circuitry, interfaces, and/or codes for executing various operations, such as user matching based on interests or the like. The processing circuitry 120 may be configured to host and enable the quiz console 116 running on (or installed on) the user device 102 to execute the operations associated with the system 100 by communicating one or more commands and/or instructions over the communication network 106. Examples of the processing circuitry 120 may include, but are not limited to, an ASIC processor, a RISC processor, a CISC processor, a FPGA, and the like. The processing circuitry 118 may be configured to perform various operations of the system 100.

The database 122 may be configured to store the logic, instructions, circuitry, interfaces, and/or codes of the processing circuitry 120 for executing various operations. The database 122 may be further configured to store therein, data associated with users registered with the system 100. The database 122 may further be configured to store training data of one or more engines of the server 104. Furthermore, the database 122 may be configured to store data in the form of documents, keywords, encoded vectors and/or questions and answers independently and/or in pairs. Some aspects of the present disclosure are intended to include and/or otherwise cover any type of the data associated with the users registered with the system 100. Examples of the database 122 may include but are not limited to, a ROM, a RAM, a flash memory, a removable storage drive, a HDD, a solid-state memory, a magnetic storage drive, a PROM, an EPROM, and/or an EEPROM. In some aspects, a set of centralized or distributed network of peripheral memory devices may be interfaced with the server 104, as an example, on a cloud server.

The communication network 106 may include suitable logic, circuitry, and interfaces that may be configured to provide a plurality of network ports and a plurality of communication channels for transmission and reception of data related to operations of various entities (such as the user device 102 and the server 104) of the system 100. Each network port may correspond to a virtual address (or a physical machine address) for transmission and reception of the communication data. For example, the virtual address may be an Internet Protocol Version 4 (IPV4) (or an IPV6 address) and the physical address may be a Media Access Control (MAC) address. The communication network 106 may be associated with an application layer for implementation of communication protocols based on one or more communication requests from the user device 102 and the server 104. The communication data may be transmitted or received, via the communication protocols. Examples of the communication

protocols may include, but are not limited to, Hypertext Transfer Protocol (HTTP), File Transfer Protocol (FTP), Simple Mail Transfer Protocol (SMTP), Domain Name System (DNS) protocol, Common Management Interface Protocol (CMIP), Transmission Control Protocol and Internet Protocol (TCP/IP), User Datagram Protocol (UDP), Long Term Evolution (LTE) communication protocols, or any combination thereof.

In some aspects of the present disclosure, the communication data may be transmitted or received via at least one communication channel of a plurality of communication channels in the communication network 106. The communication channels may include, but are not limited to, a wireless channel, a wired channel, a combination of wireless and wired channel thereof. The wireless or wired channel may be associated with a data standard which may be defined by one of a Local Area Network (LAN), a Personal Area Network (PAN), a Wireless Local Area Network (WLAN), a Wireless Sensor Network (WSN), Wireless Area Network (WAN), Wireless Wide Area Network (WWAN), a metropolitan area network (MAN), a satellite network, the Internet, a fiber optic network, a coaxial cable network, an infrared (IR) network, a radio frequency (RF) network, and a combination thereof. Aspects of the present disclosure are intended to include or otherwise cover any type of communication channel, including known, related art, and/or later developed technologies. The system 100 may further include a payment gateway 108, configured to access bank account details and/or money from bank/s of the users for payment and/or settlement of bills.

In an aspect of the present disclosure, FIG. 2 represents a block diagram that illustrates the server 104 of FIG. 1. The server 104 may include the processing circuitry 120 and the database 122. The server 104 may further include a network interface 200 and an input/output (I/O) interface 202. The processing circuitry 120, the database 122, the network interface 200, and the input/output (I/O) interface 202 may communicate with each other by way of a first communication bus 204.

In some aspects of the present disclosure, the processing circuitry 120 may include a registration engine 206, an authentication engine 208, a display engine 210, a notification generator 212, a text generation engine 214, a pre-processing engine 216, a keyword generation engine 218, a question-answer engine 220, a matching engine 222 and a grading engine 224. The server 104 may further include a second communication bus 236 to facilitate communicative coupling between one or more engines of the processing circuitry 120. It will be apparent to a person having ordinary skill in the art that the server 104 is for illustrative purposes and not limited to any specific combination of hardware circuitry and/or software.

The database 122 may include suitable logic, instructions, circuitry, interfaces, and/or codes to store a plurality of data on the server 104, associated with a registered user. The network interface 200 may include suitable logic, circuitry, and interfaces that may be configured to establish and enable a communication between the server 104 and other components of the system 100, via the communication network 106.

The network interface 200 may be implemented by use of various known technologies to support wired or wireless communication of the server 104 with the communication network 106. The network interface 200 may include, but is not limited to, an antenna, a radio frequency (RF) transceiver, one or more amplifiers, a tuner, one or more oscillators, a digital signal processor, a coder-decoder (CODEC) chipset, a subscriber identity module (SIM) card, and a local buffer circuit.

The I/O interface 202 may include suitable logic, circuitry, interfaces, and/or code that may be configured to receive inputs (e.g., orders) and transmit server outputs via a plurality of data ports in the server 104. The I/O interface 202 may include various input and output data ports for different I/O devices. Examples of such I/O devices may include, but are not limited to, a touch screen, a keyboard, a mouse, a joystick, a projector audio output, a microphone, an image-capture device, a liquid crystal display (LCD) screen and/or a speaker.

In some aspects of the present disclosure, the database 122 may include a registration data repository 226, a first repository 228, a second repository 230, a quiz memory 232 and a response memory 234. In other aspects, the database may include combined memory for one or more of the abovementioned repositories.

In some aspects, the registration data repository 226 may include registration data of the users registered on the quiz console 116. The first repository 228 may be configured to store a first set of documents and/or one or more processed forms of the first set of documents. The processed forms of the first set of documents may include but not limited to a first text, a first set of question-answer pairs and a first set of keywords. The second repository 230 may be configured to store a second set of question-answer pairs and/or one or more processed forms of the second set of question-answer pairs. The processed forms of the second set of question-answer pairs may include but not limited to a second text and a second set of keywords.

In some aspects, the quiz memory 232 may be configured to store a quiz data such that the quiz data is generated by selecting one or more of the question-answer pairs from the first repository matching with the question-answer pairs from the second repository. The response memory 234 may be

configured to store one or more question-answer pairs received from a user to be evaluated and graded by the quiz generation system 100.

In an aspect of the present disclosure, the registration engine 206 may be configured to enable the users to register into the system 100 by providing registration data through a registration menu (not shown) of the quiz console 116 displayed through the user device 102. The registration data may include, but not limited to, some personal details of the user such as name, age, verification ID, email ID and/or the like.

In some aspects of the present disclosure, the registration data may further include registration photographs of the user. In some aspects of the present disclosure, the registration engine 206 may be configured to face match the photograph taken with a phone camera (not shown) initially with a camera preview background snapshot. The registration engine 206 may be configured to utilize a facial recognition and/or a face matching algorithm to match the photographs uploaded by the user. Further, the registration engine 206 may be configured to utilize a motion verification technique such that the registration engine 206 generates random motions to be detected via the phone camera preview and verify an authenticity of the user. In another aspect, the device memory 114 may be configured to temporarily store the personal data obtained from the user by way of user interface 110, before registration and authentication of the user.

In some aspects, the authentication engine 208 may be configured to fetch personal registration data provided by the user from the device memory and authenticate the user's profile based on the comparison and verification of information provided by the user. Upon successful authentication of the user, the notification generator 212 may be configured to generate a successful authentication notification that may be displayed through the registration menu as a pop-up notification. The notification generator 212 may further be configured to generate a failed authentication notification as a pop-up upon failed authentication by the authentication engine 208. The notification may be displayed through the output interface by way of the display engine 210. In some aspects, the notification generator 212 may be configured to generate any other kind of notifications corresponding to one or more events of the system 100.

In some aspects, the text generation engine 214 may be configured to fetch the first set of documents from the user by way of the user interface 110 and generate the first text corresponding to the first set of documents. Further, the text generation engine 214 may be configured to fetch the second set of question-answer pairs from the second repository 230 and generate the second text corresponding to the second set of question-answer pairs. In some aspects, the text generation engine 214 may be based

on optical character recognition. In other aspects, the text generation engine 214 may further utilize natural language processing for generating the first text and second text.

In some aspects, the pre-processing engine 216 may be configured to pre-process the first text and the second text to generate a first string and a second string respectively. In some aspects, the pre-processing engine 216 may remove spaces, punctuation marks and stop words from the first text and the second text to generate the first string and the second string respectively.

In some aspects, the keyword generation engine 218 may be configured to divide the first string and the second string in a pre-defined length of chunks to generate the first set of keywords and the second set of keywords respectively. In some aspects, the question answer engine 220 may be configured to generate the first set of question-answer pairs corresponding to the first text.

In some aspects, the matching engine 222 may be configured to encode the first set of keywords and the second set of keywords to generate a set of first encoded vectors and the set of second encoded vectors respectively. Further, the matching engine 222 may be configured to determine a set of hamming distances between each encoded vector of the set of first encoded vectors and each encoded vector of the set of second encoded vectors. Furthermore, the matching engine 222 may be configured to match one or more encoded vectors from the set of first encoded vectors to one or more encoded vectors from the set of second encoded vectors having respective hamming distances less than a pre-defined threshold.

In some aspects, the grading engine 224 may be configured to generate grades for one or more users by matching the question-answers obtained from a user and stored in the response memory to a quiz data.

FIG. 3 represents a flow diagram that illustrates a method for quiz generation, according to an aspect of the present disclosure.

At step 302, the system 100 may be configured to receive the first set of documents from the user by way of a user interface 110 and may be stored in the first repository 228.

At step 304, the system 100 may be configured to retrieve the first set of documents and may generate the first text from the first set of documents.

At step 306, the system 100 may be configured to pre-process the first text. To pre-process the first text, the system 100 may be configured to remove spaces, punctuation marks and stop words from the first text to generate a first string.

At step 308, the system 100 may be configured to generate the first set of question-answer pairs from the first text.

At step 310, the system 100 may be configured to generate the first set of keywords from the first text. In some aspects of the present disclosure, to generate the first set of keywords, the system 100 may be configured to generate a first embedding vector, and a first string of n-gram phrases from the first text. The system 100 may further be configured to generate a first set of n-gram embedding vectors corresponding to the first string of n-gram phrases. Furthermore, the system 100 may be configured to determine a first similarity function between the first embedding vector and the first set of n-gram embedding vectors. In some aspects, based on the first similarity function, the system 100 may be configured to generate the first set of keywords.

At step 312, the system 100 may be configured to fetch the second set of question-answer pairs from the second repository.

At step 314, the system 100 may be configured to generate the second text from the second set of question-answer pairs.

At step 316, the system 100 may be configured to pre-process the second text. In some aspects, to pre-process the second text, the system 100 may be configured to pre-process the second text to generate the second string.

At step 318, the system 100 may be configured to generate the second set of keywords from the second text.

In some aspects of the present disclosure, to generate the second set of keywords, the system 100 may be configured to generate a second embedding vector, and a second string of n-gram phrases from the second text. The system 100 may further be configured to generate a second set of n-gram embedding vectors corresponding to the second string of n-gram phrases. Furthermore, the system 100 may be configured to determine a second similarity function between the second embedding vector and the second set of n-gram embedding vectors. In some aspects, based on the second similarity function, the system 100 may be configured to generate the second set of keywords.

At step 320, the system 100 may be configured to generate an encoded vector corresponding to each keyword of the first set of keywords. Further, at step 320, the system 100 may be configured to generate an encoded vector corresponding to each keyword of the second set of keywords.

In some aspects of the present disclosure, the system 100 may be configured to generate a first set of binary vectors corresponding to the first set of encoded vectors based on a pre-defined first binary threshold. The system 100 may further be configured to generate a second set of binary vectors corresponding to the second set of encoded vectors based on a pre-defined second binary threshold.

In some aspects, the pre-defined first binary threshold may be equal to the pre-defined second binary threshold.

At step 322, the system 100 may be configured to determine a hamming distance between each encoded vector corresponding to the first set of keywords and each encoded vector corresponding to the second set of keywords.

In some aspects, the system 100 may be configured to determine the hamming distance between each encoded vector corresponding to the first set of keywords and each encoded vector corresponding to the second set of keywords based on the comparison between the first set of binary vectors and the second set of binary vectors.

At step 324, the system 100 may be configured to discard the question-answer pairs from the first set of question-answer to which the hamming distance is greater than a pre-defined threshold.

At step 326, the system 100 may be configured to add the question-answer pairs having similar keywords to the second set of keywords to the quiz memory. In some aspects, the similarity is determined in terms of the hamming distance.

At step 328, the system 100 may be configured to update the second set of question-answer pairs by adding the similar questions from the first set of the question-answer pairs to the second set of question-answer pairs.

FIG. 4 represents a flow diagram that illustrates the steps in generating the first set of question-answer pairs, shown as 308 in FIG.3, according to an aspect of the present disclosure.

At step 402, the system 100 may be configured to Process the first string through a first set of encoding functions to generate a first set of vectors.

At step 404, the system 100 may be configured to process the first set of vectors through a first set of decoding functions to generate a second set of vectors.

At step 406, the system 100 may be configured to process the second set of vectors through a first linear function to generate a third set of vectors.

At step 408, the system 100 may be configured to process the third set of vectors through a first mathematical function to generate a fourth set of vectors.

At step 410, the system 100 may be configured to process the fourth set of vectors through a first tokenizer function to generate a first set of questions.

At step 412, the system 100 may be configured to Process the first string and the first set of questions through a number of encoding functions to generate a fifth set of vectors.

At step 414, the system 100 may be configured to process the fifth set of vectors through a number of decoding functions to generate a sixth set of vectors.

At step 416, the system 100 may be configured to process the sixth set of vectors through a second linear function to generate a seventh set of vectors.

At step 418, the system 100 may be configured to process the seventh set of vectors through a second mathematical function to generate an eighth set of vectors.

At step 420, the system 100 may be configured to process the eighth set of vectors through a second tokenizer function to generate a first set of answers.

At step 422, the system 100 may be configured to combine the first set of questions and the first set of answers to generate the first set of question-answer pairs.

FIG. 5 represents a flow diagram that illustrates a method 500 of generating grades by the quiz generation system, according to an aspect of the present disclosure.

At step 502, the system may be configured to retrieve the set of question-answer pairs from the quiz memory.

At step 504, the system 100 may facilitate an instructor to determine if an alteration is required in the set of the question-answer pairs.

At step 506, if there is a requirement of one or more alterations to one or more question-answer pairs of quiz data, the system 100 may facilitate the instructor to alter, change and/or remove data.

At step 508, the system 100 may be configured generate a finalized set of question answer pairs in the quiz memory.

At step 510, the system 100 may be configured to obtain a set of question-answer pairs from one or more attendee stored in the response memory 234.

At step 512, the system 100 may be configured to check similarity between the question-answer pairs from the attendees and the finalized question-answer pairs from the quiz memory.

At step 514, the system 100 may be configured to evaluate the answers from the attendees based on the similarity and generate a grade.

In some aspects, the system 100 may be configured to generate quiz including objective questions-answer pairs, subjective question-answer pairs and/or a combination thereof.

In some aspects, the keyword engine 218 may be configured to, but not limited to operate on bidirectional encoder representations from transformers (BERT) model to generate the first set of keywords and the second set of keywords.

In some aspects, the question-answer engine 220 may be configured to, but not limited to operate on a text-to-text transformer (T5 transformer) model to generate subjective and/or objective question-answer pairs.

In some aspects, the matching engine 222 may be configured to, but not limited to operate on locality sensitive hashing (LSH) model to check similarity between the first set of encoded vectors and the second set of encoded vectors for matching the first set of keywords to the second set of keywords.

As will be readily apparent to those skilled in the art, aspects of the present disclosure may easily be produced in other specific forms without departing from their essential characteristics. Aspects of the present disclosure are, therefore, to be considered as merely illustrative and not restrictive, the scope being indicated by the claims rather than the foregoing description, and all changes which come within therefore intended to be embraced therein.

As one skilled in the art will appreciate, the system includes a number of functional blocks in the form of a number of units and/or engines. The functionality of each unit and/or engine goes beyond merely finding one or more computer algorithms to carry out one or more procedures and/or methods in the form of a predefined sequential manner, rather each engine explores adding up and/or obtaining one or more objectives contributing to an overall functionality of the system. Each unit and/or engine may not be limited to an algorithmic and/or coded form, rather may be implemented by way of one or more hardware elements operating together to achieve one or more objectives contributing to the overall functionality of the system. Further, as it will be readily apparent to those skilled in the art, all the steps, methods and/or procedures of the system are generic and procedural in nature and are not specific and sequential.

Certain terms are used throughout the following description and claims to refer to particular features or components. As one skilled in the art will appreciate, different persons may refer to the same feature or component by different names. This document does not intend to distinguish between components or features that differ in name but not structure or function. While various aspects of the present disclosure have been illustrated and described, it will be clear that the present disclosure is not limited to these aspects only. Numerous modifications, changes, variations, substitutions, and equivalents will be apparent to those skilled in the art, without departing from the spirit and scope of the present disclosure, as described in the claims.

Claims:

1. A quiz generation system (100) comprising:
 - a database (122) comprising:
 - a first repository (228) configured to store a first set of documents; and
 - a second repository (230) configured to store a second set of question-answer pairs;
 - and
 - processing circuitry (120) that is coupled to the database (122) and configured to:
 - generate a first text from the first set of documents;
 - generate a first set of keywords and the first set of question-answer pairs from the first text;
 - generate a second set of keywords from the second set of question-answer pairs;
 - match the second set of keywords with the first set of keywords; and
 - select a number of question-answer pairs from the first set of question-answer pairs based on the matching of the second set of keywords with the first set of keywords.
2. The quiz generation system (100) as claimed in claim 1, wherein, to generate the first set of keywords, the processing circuitry (120) is configured to:
 - (i) pre-process the first text to generate a first string; and
 - (ii) process the first string to generate the first set of keywords.
3. The quiz generation system (100) as claimed in claim 1, wherein to generate the second set of keywords, the processing circuitry (120) is configured to:
 - (i) generate the second set of question-answer pairs to a second text;
 - (ii) pre-process the second text to generate a second string; and
 - (iii) process the second string to generate the second set of keywords.
4. The quiz generation system (100) as claimed in claim 1, wherein, to match the second set of keywords with the first set of keywords, the processing circuitry (120) is configured to:
 - (i) encode each keyword from the first set of keywords to generate a first set of encoded vectors;
 - (ii) encode each keyword from the second set of keywords to generate a second set of encoded vectors;

- (iii) determine the distance between each vector of the first set of encoded vectors and each vector of the second set of encoded vectors; and
 - (iv) match the vectors from the first set of encoded vectors to the vectors of the second set of encoded vectors having distance less than a pre-defined threshold value.
5. The quiz generation system (100) as claimed in claim 1, further comprising a quiz memory (232) that is coupled to the processing circuitry (120) and is configured to store the selected question-answer pairs from the first set of the question-answer pairs.
 6. The quiz generation system (100) as claimed in claim 1, wherein the selected question-answer pairs are iteratively updated based on the matching of the second set of keywords with the first set of keywords.
 7. The quiz generation system (100) as claimed in claim 1, further configured to generate grades by evaluating a set of answers to a set of subjective questions.
 8. A method for quiz generation (300) comprising:
 - converting a first set of documents to a first text;
 - generating a first set of keywords from the first text by pre-processing the first text to generate a first string and processing the first string to generate the first set of keywords;
 - generating a first set of question-answer pairs from the first text;
 - generating a second set of keywords from a second set of question-answer pairs by:
 - (i) converting the second set of question-answer pairs to a second text;
 - (ii) removing spaces, punctuation marks and stop words from the second text to create a second string; and
 - (iii) processing the second string to generate the second set of keywords;
 - matching the second set of keywords with the first set of keywords;
 - selecting the question-answer pairs from the first set of question-answer pairs with corresponding keywords matching to the second set of keywords.
 9. The method for quiz generation (300) as claimed in claim 8, wherein the second set of keywords are matched to the first set of keywords by:

encoding each keyword from the first set of keywords to generate a first set of encoded vectors;

encoding each keyword from the second set of keywords to generate a second set of encoded vectors;

determining the distance between each vector of the first set of encoded vectors and each vector of the second set of encoded vectors; and

matching the vectors from the first set of encoded vectors to the vectors of the second set of encoded vectors having distance less than a pre-defined threshold value.