**Title:** EXTRACTING DATA FROM COMMUNICATIONS RELATED TO DOCUMENTS

**Abstract:** The disclosed embodiments provide a system that processes data. During operation, the system obtains a communication associated with a document and extracts data associated with the document from the communication. Next, the system uses the extracted data from the communication and document data from the document to build a context associated with the document. The system then uses the context to facilitate use of the document by a user associated with the communication.
(84) Designated States (unless otherwise indicated, for every Published:  

**Mnd of regional protection available):** ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,  

EXTRACTING DATA FROM COMMUNICATIONS RELATED TO DOCUMENTS

Inventors: Sunil H. Madhani, Joseph J. O'Sullivan and Anu Sreepathy

BACKGROUND

[0001] The disclosed embodiments relate to techniques for extracting data. More specifically, the disclosed embodiments relate to techniques for extracting data from communications related to documents.

[0002] Data processing and exchange are essential to many business and personal transactions. For example, small businesses may use accounting and/or inventory data to obtain and share reports regarding inventory sales, customer invoices, and/or cash flow. Similarly, healthcare providers may examine medical records to view patient information related to insurance providers, medical conditions, and/or office visits.

[0003] In addition, data exchange among users frequently involves the use of documents such as word-processing documents, spreadsheets, and/or Portable Document Format (PDF) documents. For example, a business may manage business transactions with a set of customers by creating a set of bills, invoices, and/or other types of documents containing data associated with the business transactions and transmitting the documents to the respective customers via email. The customers may use the data in the documents to pay the bills and/or invoices, respond to the business, and/or update their records of the transactions.

[0004] However, information relevant to document-based transactions is frequently found outside of the documents. As a result, a user that relies solely on information from a document to perform a transaction related to the document may not be aware of other options for performing the transaction, which are specified and/or described outside of the document. For example, a business may send an invoice as an attachment in an email to a customer. The business may also include discounts or promotions related to products or services in the invoice and relevant dates (e.g., deadlines, periods of validity, etc.) for the discounts or promotions in the body of the email. If the customer ignores the email's content and pays the invoice based solely on the amounts and
due dates in the invoice, the customer may pay a higher amount than if the customer took advantage of the discounts or promotions in the email.

[0005] Consequently, use of documents may be facilitated by mechanisms for supplementing data in the documents with relevant data that is found outside of the documents.

SUMMARY

[0006] The disclosed embodiments provide a system that processes data. During operation, the system obtains a communication associated with a document and extracts data associated with the document from the communication. Next, the system uses the extracted data from the communication and document data from the document to build a context associated with the document. The system then uses the context to facilitate use of the document by a user associated with the communication.

[0007] In some embodiments, the system also uses information from the communication to obtain an additional communication related to the communication, and includes data from the additional communication in the extracted data.

[0008] In some embodiments, extracting the data from the communication involves applying a domain-specific grammar to the communication to obtain the data from the communication.

[0009] In some embodiments, using the extracted data and the document data to build the context associated with the communication involves applying the extracted data to the document data to obtain one or more facts and one or more insights, and including the one or more facts and the one or more insights in the context.

[0010] In some embodiments, using the context to facilitate use of the document by the user involves at least one of:

(i) providing the context to the user;
(ii) generating a reminder associated with the context;
(iii) recommending a user action based on the context; and
(iv) performing an action for the user based on the context.

[0011] In some embodiments, the document includes a financial document.

[0012] In some embodiments, the extracted data includes at least one of a discount, a date, and a promotion.

[0013] In some embodiments, the communication is at least one of an email, a text message, an instant message, a social-networking post, and a voicemail.
In some embodiments, the document is provided using an attachment or a link in the communication.

BRIEF DESCRIPTION OF THE FIGURES

[0015] FIG. 1 shows a schematic of a system in accordance with the disclosed embodiments.

[0016] FIG. 2 shows the creation of a context associated with a document in accordance with the disclosed embodiments.

[0017] FIG. 3 shows a flowchart illustrating the processing of data in accordance with the disclosed embodiments.

[0018] FIG. 4 shows a computer system in accordance with the disclosed embodiments.

[0019] In the figures, like reference numerals refer to the same figure elements.

DETAILED DESCRIPTION

The following description is presented to enable any person skilled in the art to make and use the embodiments, and is provided in the context of a particular application and its requirements. Various modifications to the disclosed embodiments will be readily apparent to those skilled in the art, and the general principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the present disclosure. Thus, the present invention is not limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles and features disclosed herein.

[0021] The disclosed embodiments provide a method and system for processing data. The data may include document data from a document such as a bill, invoice, form, statement, transcript, record, receipt, and/or tax document. The data may also include data extracted from one or more communications related to the document. For example, the data may be obtained from emails, text messages, instant messages, and/or voicemails and include discounts, promotions, additional terms, verification data, and/or other options related to a transaction associated with the document. As a result, the data may be applied to and/or combined with the document data to improve use of the document data and/or carrying out of a transaction associated with the document by a user.

[0022] As shown in FIG. 1, the data may be included in a document 110 such as a word-processing document, spreadsheet, Portable Document Format (PDF) document, web page, and/or another type of digitized document. For example, document 110 may be a bill, invoice, form, statement, transcript, record, receipt, tax document, and/or other type of document that
includes a specific type and/or set of data. Document 110 may be generated digitally using a
word-processing application, spreadsheet application, diagramming application, graphics-editing
application, and/or other application for creating electronic documents. Alternatively, document
110 may correspond to a scanned and/or photographed digital copy of a physical document.

[0023] In one or more embodiments, the system of FIG. 1 includes functionality to
automatically extract data from document 110 without requiring use of a template and/or custom
code for extracting the data. As shown in FIG. 1, the system may correspond to a data-
processing framework 102 that provides a document-processing apparatus 106 for obtaining
document data 114 from document 110. For example, document-processing apparatus 106 may
analyze the contexts of individual words in document 110 and store data associated with the
words and contexts in a set of data elements and tags. As a result, document-processing
apparatus 106 may be capable of obtaining document data 114 from document 110 in the absence
of a template that specifies the locations of data elements in document 110. Alternatively, if a
template for document 110 is available, document-processing apparatus 106 may use the
template to extract document data 114 from document 110.

[0024] Document data 114 from document 110 may then be used in a transaction
involving two or more users. For example, a business owner may provide an invoice to a
customer as an attachment in an email. The customer may download the attachment, open the
invoice, and use information in the invoice to pay the business owner for goods or services
specified in the invoice. Moreover, if document data 114 has been extracted from the invoice by
document-processing apparatus 106, document data 114 may be provided to a management
apparatus 130 in data-processing framework 102 and/or a financial-management application
associated with data-processing framework 102. Management apparatus 130 and/or the
application may use document data 114 to generate one or more payment reminders for the
invoice and/or schedule a payment of the invoice for the customer.

[0025] Those skilled in the art will appreciate that additional information related to
document 110 may be found in one or more communications 116-118 associated with document
110. The additional information may facilitate use of document 110 and/or carrying out of a
transaction associated with document 110. Continuing with the above example, the business
owner may specify additional terms associated with the invoice in the body of the email. The
business owner may also inform the customer of discounts or promotions related to the invoice in
the same email and/or in a separate email to the customer. Additional discounts, promotions,
and/or options associated with payment of the invoice may further be found in emails from other
senders, such as "daily deal" websites and/or affiliates. Thus, if the customer pays the invoice
based solely on the amounts and/or due dates in the invoice, the customer may pay a higher amount than if the customer took advantage of the discounts or promotions in the email(s).

[0026] In one or more embodiments, data-processing framework 102 includes functionality to enhance and/or supplement document data 114 with additional data 112 extracted from one or more communications 116-118 associated with document 110. In turn, the additional data 112 may facilitate use of document 110 and/or participation in a transaction associated with document 110 by a user associated with document 110.

[0027] More specifically, a communication-processing apparatus 104 may obtain communication 116 associated with document 110 and extract data 112 associated with document 110 from communication 116. Communication 116 may include an email, text message, instant message, voicemail, and/or other form of digital and/or digitized communication involving two or more users. Communication 116 may be stored in a communication repository 120 with one or more additional communications 118. For example, communications 116-118 may be obtained from an email inbox, a text or instant message thread, a voicemail inbox, cloud storage, and/or a backup mechanism. A crawler associated with communication-processing apparatus 104 may detect the arrival of new communications in communication repository 120, or the crawler and/or communication-processing apparatus 104 may periodically obtain and process batches of communications from communication repository 120.

[0028] Communication-processing apparatus 104 may also use communication 116 to obtain document 110. For example, communication-processing apparatus 104 may obtain document 110 as an attachment to communication 116. Alternatively, communication-processing apparatus 104 may access document 110 through a link and/or reference to document 110 in communication 116. Communication-processing apparatus 104 may provide document 110 to document-processing apparatus 106 for extraction of document data 114 from document 110, as described above and in the above-referenced application.

[0029] To extract data 112 from communication 116, communication-processing apparatus 104 may scan communication 116 to obtain a sender, recipient, and/or document 110 from communication 116. Communication-processing apparatus 104 may also detect a domain related to communication 116 based on one or more keywords found in communication 116. For example, communication-processing apparatus 104 may use a natural-language-processing (NLP) technique to parse text in communication 116 and identify the sender, recipient, and/or document 110 associated with communication 116. Communication-processing apparatus 104 may also match keywords in the parsed sentences to one or more domains (e.g., application
domains, business domains, etc.) supported by data-processing framework 102. If communication 116 is not text-based (e.g., voicemail, video chat, phone call, etc.), communication-processing apparatus 104 and/or another component of data-processing framework 102 may use a speech-recognition technique and/or another mechanism to convert communication 116 into text before scanning communication 116.

[0030] Next, communication-processing apparatus 104 may apply a domain-specific grammar (e.g., domain-specific grammar 112, domain-specific grammar 1 126, domain-specific grammar 2 128) to communication 116 to obtain domain-specific words and/or data 112 from communication 116. For example, communication-processing apparatus 104 may obtain one or more domain-specific grammars for domains matched by keywords in communication 116 from a grammar repository 124. Communication-processing apparatus 104 may use the domain-specific grammar(s) to detect structures in communication 116 represented by the domain-specific grammar(s) and extract data elements from the structures, as described in further detail below with respect to FIG. 2.

[0031] Communication-processing apparatus 104 may additionally use data 112 and/or other information from communication 116 to identify and/or obtain one or more additional communications 118 related to communication 116. For example, communication-processing apparatus 104 may obtain additional communications 118 associated with the same sender and/or domain as communication 116. As with communication 116, communication-processing apparatus 104 may parse communications 118 and/or apply domain-specific grammars to communications 118 to obtain additional data from communications 118. Communication-processing apparatus 104 may then supplement data 112 with the additional data.

[0032] After data 112 is extracted from communication 116 and/or related communications 118, a context-building apparatus 108 in data-processing framework 102 may build a context 122 associated with document 110 using data 112 and document data 114 from document-processing apparatus 106. In particular, context-building apparatus 108 may apply data 112 to document data 114 to obtain context 122 as one or more facts and/or one or more insights. For example, context-building apparatus 108 may apply a discount found in communication 116 and/or communications 118 to an amount due in an invoice to obtain a discounted amount for the invoice and/or a pay-by date for the discounted amount.

[0033] Finally, a management apparatus 130 in data-processing framework 102 may use context 122 to facilitate use of document 110 by a user associated with communication 116. Management apparatus 130 may provide facts and/or insights from context 122 to the user. For example, management apparatus 130 may inform the user of the facts and/or insights from
context-building apparatus 108 in one or more notifications. Management apparatus 130 may also generate a reminder associated with context 122. For example, management apparatus 130 may remind the user of a due date and/or deadline associated with document 110 prior to the due date and/or deadline. Finally, management apparatus 130 may recommend a user action and/or perform an action for the user based on context 122. For example, management apparatus 130 may recommend payment of a bill or invoice during a period associated with a discount on the bill or invoice, or management apparatus 130 may automatically schedule a payment to occur within the period to ensure that the user receives the discount.

[0034] By supplementing document data (e.g., document data 114) from documents (e.g., document 110) with data (e.g., data 112) extracted from communications (e.g., communications 116-118) related to the documents, data-processing framework 102 may enable better understanding and use of information in the documents by users associated with the documents. In turn, data-processing framework 102 may optimize subsequent communications and/or transactions associated with the documents for the users, thus reducing overhead and/or inefficient handling of the transactions by the users.

[0035] Those skilled in the art will appreciate that the system of FIG. 1 may be implemented in a variety of ways. First, communication-processing apparatus 104, document-processing apparatus 106, context-building apparatus 108, management apparatus 130, and grammar repository 124 may be provided by a single physical machine, multiple computer systems, a grid, one or more databases, one or more filesystems, and/or a cloud computing system. In addition, communication-processing apparatus 104, document-processing apparatus 106, context-building apparatus 108, and/or management apparatus 130 may be provided by the same hardware and/or software component or execute independently from one another. For example, communication-processing apparatus 104, document-processing apparatus 106, context-building apparatus 108, and management apparatus 130 may form a part of an application that is used to process and manage data from a set of documents and related communications for a user. On the other hand, one or more components of data-processing framework 102 may execute outside the application and provide data from the documents and/or communications in a form that can be used by the application and/or other applications after the data is extracted from the documents and/or communications.

[0036] Second, data-processing framework 102 may be configured for use with various types of electronic documents, communications, and/or data. For example, data-processing framework 102 may accept multiple types and/or file formats of documents and/or include functionality to convert the documents into a file format that is suitable for text and/or data
extraction. Along the same lines, data-processing framework 102 may be capable of extracting data from various types of communications, including text-based communications (e.g., emails, text messages, instant messages, social-networking posts, forum messages, etc.) and/or other types of communications (e.g., phone calls, video chats, voicemails, etc.). Data-processing framework 102 may further obtain and/or store data 112, document data 114, and/or domain-specific grammars as relational-database records, property lists, Extensible Markup Language (XML) documents, JavaScript Object Notation (JSON) objects, and/or other types of structured data.

[0037] FIG. 2 shows the creation of a context 212 associated with a document 206 in accordance with the disclosed embodiments. As mentioned above, document 206 may be obtained from a communication 202 such as an email, text message, instant message, social-networking post, and/or voicemail. For example, document 206 may be included in communication 202 as an attachment, or document 206 may be obtained using a link and/or reference in communication 202. Document data 210 may also be extracted from document 206, with or without a template for document 206.

[0038] Next, communication 202 may be scanned to obtain information such as a sender, recipient, and/or domain 218 associated with communication 202. For example, an NLP technique may be used to identify the sender, recipient, and/or other entities associated with an email containing an invoice. The NLP technique may also match one or more keywords (e.g., "invoice," "pay," "payment," "business," etc.) in the email and/or invoice to a domain related to invoice processing and/or financial transactions.

[0039] A domain-specific grammar 204 for domain 218 may also be applied to communication 202 to obtain data 208 associated with document 206. For example, a domain-specific grammar for invoice processing may include the following:

```
INVOICE [
  DISCOUNT*
  DATE:d
  AMOUNT:a
  DISTANCED
]
```

(Avail x% discount if paid by d)

(Get a discount of a <filler>)

(Obtain a [rebate |discount] of [x%|a] as <filler>) (x% off if paid via <filler>)
(x% waived <filler> shipping is within r miles)
(No shipping cost<filler> for order [more |greater |higher] than a)
]
As a result, the domain-specific grammar may specify one or more structures that allow data 208
(e.g., discount, date, amount, distance) related to discounts and/or promotions for an invoice to
be extracted from text in communication 202. The domain-specific grammar 204 may then be
applied to a sentence such as "Avail 5% discount if paid before 11/20/2013" to obtain a discount
of 5% and a date of "11/20/2013" from the sentence.

[0040] Domain 218, data 208, and/or other information from communication 202 may
further be used to obtain an additional communication 220 related to communication 202.
Domain-specific grammar 204 and/or another domain-specific grammar may be applied to
communication 220, and data extracted from communication 220 may be included in data 208.
If more communications related to communication 202 are found, the domain-specific
grammar(s) may also be applied to the communications, and data extracted from the
communications may be added to data 208.

[0041] Data 208 extracted from communication 202 and/or communication 220 and
document data 210 from document 206 may then be used to build context 212. In particular, data
208 may be applied to document data 210 to obtain one or more facts 214 and one or more
insights 216 that are included in context 212. For example, document data 210 for an invoice
may include the following:

Vendor name - John's Furnishings
Due Date - 11/30/2013
Bill Date - 10/20/2013
Amount Due - $546.08

Line Item 1 - Window blinds $380.00
Line Item 2 - Cushions $166.08

Similarly, data 208 from communication 202 and/or communication 220 may include a discount
of 10% and a due date of November 20, 2013. Data 208 may then be applied to document data
210 to obtain an insight containing a discounted amount of $491.48 and a pay-by date for the
discounted amount of November 20, 2013. Data 208, document data 210, and/or the insight may
then be provided to a user in the following form:

"$546.08 due to John's Furnishings by 11/30/2013. Pay only $491.48 if paid by
11/20/2013."

Consequently, context 212 may include additional information that is relevant to the invoice but not present in the invoice. The additional information may allow the user to take advantage of a discount that the user may otherwise be unaware of, thus facilitating use and/or payment of the invoice by the user.

FIG. 3 shows a flowchart illustrating the processing of data in accordance with the disclosed embodiments. In one or more embodiments, one or more of the steps may be omitted, repeated, and/or performed in a different order. Accordingly, the specific arrangement of steps shown in FIG. 3 should not be construed as limiting the scope of the embodiments.

Initially, a communication associated with a document is obtained (operation 302). The communication may be an email, a text message, an instant message, a social-networking post, a voicemail, and/or another type of digital or digitized communication. Next, data associated with the document is extracted from the communication (operation 304). For example, a domain-specific grammar may be applied to the communication to identify structures in the communication and obtain data from the structures.

An additional communication related to the communication may be available (operation 306). For example, the additional communication may be obtained using information (e.g., sender, recipient, domain, keywords, etc.) and/or data from the communication. If an additional communication is available, data from the additional communication is included in the extracted data (operation 308). For example, the domain-specific grammar may be applied to the additional communication to obtain additional data related to the communication and/or document. If no additional communication is available, no additional data is included in the data extracted from the communication.

The extracted data and document data from the document are then used to build a context associated with the document (operation 310). For example, the extracted data may be applied to the document data to obtain one or more facts and one or more insights, and the fact(s) and/or insight(s) may be included in the context.

Finally, the context is used to facilitate use of the document by a user associated with the communication (operation 312), such as a recipient of the communication. To facilitate use of the document by the user, the context may be provided to the user to notify the user of the facts and/or insights obtained from the extracted data and/or document data. For example, the user may be notified of available discounts or promotions, errors or discrepancies, and/or changes in terms or additional terms associated with the document. A reminder associated with the context may also be generated to allow the user to take advantage of the insights in a timely manner. For example, a reminder may be generated to remind the user of a pay-by date for a
discount or promotion. Finally, the context may be used to recommend a user action and/or perform an action for the user. For example, the user may be provided with a phone number and/or email address for contacting another user associated with the document or communication if an error or discrepancy is found in the document or communication. Alternatively, an email requesting review or explanation of the error or discrepancy may be automatically generated and sent to the other user on the user's behalf.

[0047] FIG. 4 shows a computer system 400 in accordance with the disclosed embodiments. Computer system 400 includes a processor 402, memory 404, storage 406, and/or other components found in electronic computing devices. Processor 402 may support parallel processing and/or multi-threaded operation with other processors in computer system 400. Computer system 400 may also include input/output (I/O) devices such as a keyboard 408, a mouse 410, and a display 412.

[0048] Computer system 400 may include functionality to execute various components of the present embodiments. In particular, computer system 400 may include an operating system (not shown) that coordinates the use of hardware and software resources on computer system 400, as well as one or more applications that perform specialized tasks for the user. To perform tasks for the user, applications may obtain the use of hardware resources on computer system 400 from the operating system, as well as interact with the user through a hardware and/or software framework provided by the operating system.

[0049] In one or more embodiments, computer system 400 provides a system for processing data. The system may include a communication-processing apparatus that obtains a communication associated with a document and extracts data associated with the document from the communication. The system may also include a context-building apparatus that uses the extracted data from the communication and document data from the document to build a context associated with the document. Finally, the system may include a management apparatus that uses the context to facilitate use of the document by a user associated with the communication.

[0050] In addition, one or more components of computer system 400 may be remotely located and connected to the other components over a network. Portions of the present embodiments (e.g., communication-processing apparatus, context-building apparatus, management apparatus, etc.) may also be located on different nodes of a distributed system that implements the embodiments. For example, the present embodiments may be implemented using a cloud computing system that extracts data from communications related to a set of documents and uses the extracted data to facilitate use of the documents by a set of remote users.
The data structures and code described in this detailed description are typically stored on a computer-readable storage medium, which may be any device or medium that can store code and/or data for use by a computer system. The computer-readable storage medium includes, but is not limited to, volatile memory, non-volatile memory, magnetic and optical storage devices such as disk drives, magnetic tape, CDs (compact discs), DVDs (digital versatile discs or digital video discs), or other media capable of storing code and/or data now known or later developed.

The methods and processes described in the detailed description section can be embodied as code and/or data, which can be stored in a computer-readable storage medium as described above. When a computer system reads and executes the code and/or data stored on the computer-readable storage medium, the computer system performs the methods and processes embodied as data structures and code and stored within the computer-readable storage medium.

Furthermore, methods and processes described herein can be included in hardware modules or apparatus. These modules or apparatus may include, but are not limited to, an application-specific integrated circuit (ASIC) chip, a field-programmable gate array (FPGA), a dedicated or shared processor that executes a particular software module or a piece of code at a particular time, and/or other programmable-logic devices now known or later developed. When the hardware modules or apparatus are activated, they perform the methods and processes included within them.

The foregoing descriptions of various embodiments have been presented only for purposes of illustration and description. They are not intended to be exhaustive or to limit the present invention to the forms disclosed. Accordingly, many modifications and variations will be apparent to practitioners skilled in the art. Additionally, the above disclosure is not intended to limit the present invention.
What I s Claimed Is:

1. A computer-implemented method for processing data, comprising:
   obtaining a communication associated with a document;
   extracting data associated with the document from the communication;
   using the extracted data from the communication and document data from the document to build a context associated with the document; and
   using the context to facilitate use of the document by a user associated with the communication.

2. The computer-implemented method of claim 1, further comprising:
   using information from the communication to obtain an additional communication related to the communication; and
   including data from the additional communication in the extracted data.

3. The computer-implemented method of claim 1, wherein extracting the data from the communication involves:
   applying a domain-specific grammar to the communication to obtain the data from the communication.

4. The computer-implemented method of claim 1, wherein using the extracted data and the document data to build the context associated with the communication involves:
   applying the extracted data to the document data to obtain one or more facts and one or more insights; and
   including the one or more facts and the one or more insights in the context.

5. The computer-implemented method of claim 1, wherein using the context to facilitate use of the document by the user involves at least one of:
   providing the context to the user;
   generating a reminder associated with the context;
   recommending a user action based on the context; and
   performing an action for the user based on the context.
6. The computer-implemented method of claim 1, wherein the document comprises a financial document.

7. The computer-implemented method of claim 5, wherein the extracted data comprises at least one of a discount, a date, and a promotion.

8. The computer-implemented method of claim 1, wherein the communication is at least one of an email, a text message, an instant message, a social-networking post, and a voicemail.

9. The computer-implemented method of claim 1, wherein the document is provided using an attachment or a link in the communication.

10. A system for processing data, comprising:
    a communication-processing apparatus configured to:
    obtain a communication associated with a document; and
    extract data associated with the document from the communication;
    a context-building apparatus configured to use the extracted data from the communication and document data from the document to build a context associated with the document; and
    a management apparatus configured to use the context to facilitate use of the document by a user associated with the communication.

11. The system of claim 10, further comprising:
    a document-processing apparatus configured to extract the document data from the document.

12. The system of claim 10, wherein the communication-processing apparatus is further configured to:
    obtain an additional communication related to the communication; and
    include data from the additional communication in the extracted data.

13. The system of claim 10, wherein extracting the data from the communication involves:
applying a domain-specific grammar to the communication to obtain the data from the communication.

14. The system of claim 10, wherein using the extracted data and the document data to build the context associated with the communication involves:
   applying the extracted data to the document data to obtain one or more facts and one or more insights; and
   including the one or more facts and the one or more insights in the context.

15. The system of claim 10, wherein using the context to facilitate use of the document by the user involves at least one of:
   providing the context to the user;
   generating a reminder associated with the context;
   recommending a user action based on the context; and
   performing an action for the user based on the context.

16. A non-transitory computer-readable storage medium storing instructions that when executed by a computer cause the computer to perform a method for processing data, the method comprising:
   obtaining a communication associated with a document;
   extracting data associated with the document from the communication;
   using the extracted data from the communication and document data from the document to build a context associated with the document; and
   using the context to facilitate use of the document by a user associated with the communication.

17. The non-transitory computer-readable storage medium of claim 16, the method further comprising:
   using information from the communication to obtain an additional communication related to the communication; and
   including data from the additional communication in the extracted data.

18. The non-transitory computer-readable storage medium of claim 16, wherein extracting the data from the communication involves:
applying a domain-specific grammar to the communication to obtain the data from the communication.

19. The non-transitory computer-readable storage medium of claim 16, wherein using the extracted data and the document data to build the context associated with the communication involves:
   applying the extracted data to the document data to obtain one or more facts and one or more insights; and
   including the one or more facts and the one or more insights in the context.

20. The non-transitory computer-readable storage medium of claim 16, wherein using the context to facilitate use of the document by the user involves at least one of:
   - providing the context to the user;
   - generating a reminder associated with the context;
   - recommending a user action based on the context; and
   - performing an action for the user based on the context.
Start

Obtain communication associated with document 302

Extract data associated with document from communication 304

Additional communication available? 306

Yes

Include data from additional communication in extracted data 308

No

Use extracted data and document data from document to build context associated with document 310

Use context to facilitate use of document by user associated with communication 312

End

FIG. 3
FIG. 4
B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F 17/00; G06F 17/30; G06F 7/00; G07B 17/00; G06F 17/27; G06Q 10/00; G06F 17/21

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS/KIPO internal & Keywords: communication, email, financial document, context, attachment, domain-specific, discount, and similar terms.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>US 2012-0078610 A1 (SAVAL, MEHMET OGUZ et al.) 29 March 2012</td>
<td>1-20</td>
</tr>
<tr>
<td></td>
<td>See paragraphs [0006H0008], [0016], [0027], [0032H0033], [0035], [0041H0042]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and [0048]; claims 1-2 and 8; and figures 1-2.</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>US 2012-0030971 A1 (THOMAS, JAMES et al.) 27 December 2012</td>
<td>1-20</td>
</tr>
<tr>
<td></td>
<td>See paragraphs [0021], [0027]-[0044], [0049]-[0054], [0059], and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0081H0090]; claim 1, and figures 1 and 3.</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>US 2010-0294929 A1 (RAGGER, LOTHAR) 25 November 2010</td>
<td>1-20</td>
</tr>
<tr>
<td></td>
<td>See paragraphs [0023]-[0024] and figure 4.</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>US 7,072,061 B2 (BLAIR, WILLIAM et al.) 04 July 2006</td>
<td>1-20</td>
</tr>
<tr>
<td></td>
<td>See column 2, lines 13-18; column 6, lines 7-50;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>column 7, line 40 - column 8, line 43; and figure 3.</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>US 7,603,300 B2 (HAFFNER, PETER et al.) 13 October 2009</td>
<td>1-20</td>
</tr>
<tr>
<td></td>
<td>See column 8, line 62 - column 10, line 47; and figure 4.</td>
<td></td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:
  "A" document defining the general state of the art which is not considered to be of particular relevance
  "E" earlier application or patent but published on or after the international filing date
  "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  "O" document referring to an oral disclosure, use, exhibition or other means
  "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
17 December 2014 (17.12.2014)

Date of mailing of the international search report
17 December 2014 (17.12.2014)

Name and mailing address of the ISA/KR
International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea
Facsimile No. +82-42-472-7140

Authorized officer
NHO, Ji Myong
Telephone No. +82-42-481-8528
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family number</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>US 2012-0078610 Al</td>
<td>29/03/2012</td>
<td>US 8719007 B2</td>
<td>06/05/2014</td>
</tr>
<tr>
<td>US 2012-0330971 Al</td>
<td>27/12/2012</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>US 2010-0299229 Al</td>
<td>25/11/2010</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2003-0061316 Al</td>
<td>27/03/2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2008-0059612 Al</td>
<td>06/03/2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2008-0104097 Al</td>
<td>01/05/2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2008-0109465 Al</td>
<td>08/05/2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2008-0154930 Al</td>
<td>26/06/2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 7084998 B2</td>
<td>01/08/2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 7277878 B2</td>
<td>02/10/2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 7870115 B2</td>
<td>11/01/2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 2003-294713 Al</td>
<td>15/06/2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1563429 Al</td>
<td>17/08/2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1563434 Al</td>
<td>17/08/2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2004-0098663 Al</td>
<td>20/05/2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2004-0107123 Al</td>
<td>03/06/2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 7584192 B2</td>
<td>01/09/2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2004-046977 A2</td>
<td>03/06/2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2004-046978 A2</td>
<td>03/06/2004</td>
</tr>
</tbody>
</table>