

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2024年10月3日(03.10.2024)



(10) 国際公開番号

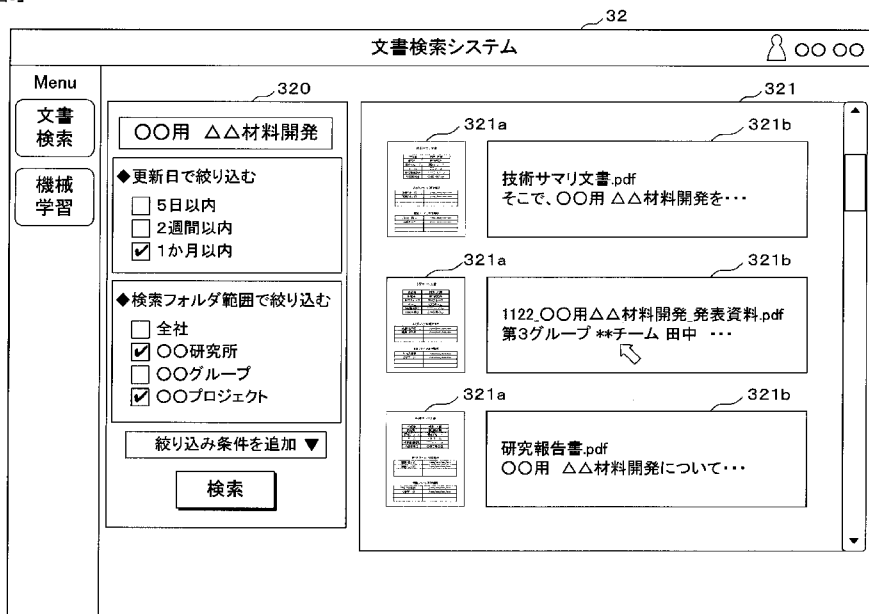
WO 2024/204142 A1

- (51) 国際特許分類:
G06F 16/34 (2019.01) G06N 20/00 (2019.01)
- (21) 国際出願番号: PCT/JP2024/011851
- (22) 国際出願日: 2024年3月26日(26.03.2024)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2023-058077 2023年3月31日(31.03.2023) JP
- (71) 出願人: 住友化学株式会社
(SUMITOMO CHEMICAL COMPANY, LIMITED) [JP/JP]; 〒1036020 東京都中央区日本橋二丁目7番1号 Tokyo (JP).
- (72) 発明者: 押木 淳(OSHIKI, Jun); 〒1036020 東京都中央区日本橋二丁目7番1号 住友化学株式会社内 Tokyo (JP). 西田 理彦(NISHIDA, Toshihiko); 〒1036020 東京都中央区日本橋二丁目7番1号 住友化学株式会社内 Tokyo (JP). 栗田 靖之(KURITA, Yasuyuki); 〒1036020 東京都中央区日本橋二丁目7番1号 住友化学株式会社内 Tokyo (JP).
- (74) 代理人: 河野 英仁, 外(KOHNO, Hideto et al.); 〒5400035 大阪府大阪市中央区釣鐘町二丁目4番3号 河野特許事務所 Osaka (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN,

(54) Title: COMPUTER PROGRAM, INFORMATION PROCESSING METHOD, AND INFORMATION PROCESSING DEVICE

(54) 発明の名称: コンピュータプログラム、情報処理方法及び情報処理装置

[図8]



(57) Abstract: The present invention provides a computer program or the like which is capable of improving the convenience of data sharing. The computer program causes a computer to execute processing which involves storing a plurality of documents created by a plurality of users in a group, outputting list information which selectably displays summary information expressing a summary of each stored document in a list, obtaining a selection of any one unit of summary information from among the plurality of units of summary information in the outputted list information, extracting detail

CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC,
EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR,
HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

一 國際調查報告（条約第21条(3)）

information expressing the details of a document corresponding to the selected summary information, and outputting the extracted detail information.

(57) 要約：データ共有の利便性を向上することができるコンピュータプログラム等を提供する。コンピュータプログラムは、グループ内における複数のユーザにより作成された複数の文書を記憶しておき、記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出力し、出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取得し、選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、抽出した前記詳細情報を出力する処理をコンピュータに実行させる。

明 細 書

発明の名称：

コンピュータプログラム、情報処理方法及び情報処理装置

技術分野

[0001] 本発明は、コンピュータプログラム、情報処理方法及び情報処理装置に関する。

背景技術

[0002] 従来、データの共有に関する技術が提案されている。例えば特許文献1には、新たな患者データ共有無しの多中心生物医学データコラボレーション処理システムであって、各医療センターモデルパラメータと非同期計算をコラボレートするためのクラウドサーバ及びデータにローカル高性能計算を展開するための医療センタークライアントを含むシステムが開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特許第7074306号公報

発明の概要

発明が解決しようとする課題

[0004] しかしながら、特許文献1に記載の技術は、必ずしもデータ共有の利便性を向上するものではないという問題がある。

[0005] 本開示の目的は、データ共有の利便性を向上することができるコンピュータプログラム等を提供することである。

課題を解決するための手段

[0006] 本開示の一態様に係るコンピュータプログラムは、グループ内における複数のユーザにより作成された複数の文書を記憶しておき、記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出力し、出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取

得し、選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、抽出した前記詳細情報を出力する処理をコンピュータに実行させる。

[0007] 本開示の一態様に係るコンピュータプログラムは、グループ内における複数のユーザにより作成された複数の文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を取得し、取得した前記一覧情報を表示し、表示した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を受け付け、受け付けた前記概要情報の選択を出力し、選択された概要情報に対応する文書の詳細を表す詳細情報を取得し、取得した前記詳細情報を表示する処理をコンピュータに実行させる。

発明の効果

[0008] 本開示によれば、データ共有の利便性を向上することができる。

図面の簡単な説明

[0009] [図1]本実施形態の検索システムの構成例を示すブロック図である。
[図2]文書DBに記憶される情報の内容例を示す図である。
[図3]表紙の内容例を示す図である。
[図4]まとめファイルの内容例を示す図である。
[図5]検索システムが実行する文書の取得処理手順の一例を示すフローチャートである。
[図6]検索システムが実行する検索処理手順の一例を示すフローチャートである。
[図7]端末装置の表示部に表示される受付画面の一例を示す模式図である。
[図8]端末装置の表示部に表示される概要画面の一例を示す模式図である。
[図9]第2実施形態の検索システムの構成例を示すブロック図である。
[図10]学習モデルの概要を示す説明図である。
[図11]第2実施形態の検索システムが実行する処理手順の一例を示すフローチャートである。

発明を実施するための形態

[0010] 本開示をその実施の形態を示す図面を参照して具体的に説明する。

[0011] (第1実施形態)

図1は、本実施形態の検索システム100の構成例を示すブロック図である。本実施形態の検索システム100は、企業（会社）内で収集された複数の文書を共有データとして管理し、ユーザに対して所望の文書を提供する。企業は、複数のユーザが属するグループの一例である。本明細書において、「グループ」とは、データの共有を行う集団を意味する。グループは、1つの企業（組織）であってもよく、複数の企業からなるものであってもよく、企業を区切るより小さな単位であってもよい。検索システム100は、情報処理装置1と、複数の端末装置2とを備える。情報処理装置1と、端末装置2それぞれとは、例えばLAN（Local Area Network）又はインターネット等のネットワークNを介して通信可能に接続されている。

[0012] 情報処理装置1は、種々の情報処理、情報の送受信が可能な装置であり、例えばサーバコンピュータ、パーソナルコンピュータ、量子コンピュータ等である。情報処理装置1は端末装置2と同じ施設に設置されたローカルコンピュータであってもよく、インターネット等を介して端末装置2に通信接続されたクラウドコンピュータであってもよい。情報処理装置1は、いずれかの端末装置2に統合されてもよい。端末装置2の数は1、2又は4以上であってもよい。

[0013] 端末装置2は、検索システム100を利用するユーザに用いられる情報処理端末である。端末装置2は、例えばパーソナルコンピュータ、スマートフォン、タブレット端末等である。以下の説明では、便宜上、第1ユーザが本システムに文書を提供するユーザであり、第2ユーザが提供された文書を検索、利用するユーザであるものとして、第1ユーザの端末装置2を端末装置2a、第2ユーザの端末装置2を端末装置2bと記載する。なお、第1ユーザ及び第2ユーザを区別して説明する必要がない場合には、単にユーザとも記載し、端末装置2a、2bを区別して説明する必要がない場合には、単に端末装置2とも記載する。第1ユーザ及び第2ユーザは、例えば同じ会社内の別チーム所属といったように、同一組織の異なる部署に属する。

[0014] 上記のユーザの区別は便宜的なものであり、文書を提供するユーザが文書を検索するユーザとなる場合も想定される。すなわち、本システムはユーザ同士の文書の共有を実現するものであり、ユーザは文書の提供者又は提供を受ける者のどちらか一方のみに分類されるものではない。

[0015] 情報処理装置 1 は、端末装置 2 a を通じて、第 1 ユーザそれぞれにより作成された文書を受け付け、受け付けた文書を後述するデータベースに保存する。データベースに保存される文書とは、例えば技術文書であり、具体的には月報、研究報告書、実験報告書、学術論文等を含む。情報処理装置 1 は、第 2 ユーザからの要求に応じて、保存された文書に関する情報を端末装置 2 b に提供する。

[0016] 従来ユーザ毎に個別に保持されていた実験結果や研究成果を、部署や業務を跨いで大量に収集し共有データとして利活用することにより、技術開発の促進効果が期待される。しかしながら、業務データの作成は業務知識に基づく属人性が高く、作成者に応じてデータ形式やデータ内容が異なるため、データの利活用を向上するためには、データ共有に適した態様でデータを保存することが重要である。特に化学分野においては、研究開発が属人化しており、実験データの取り方がユーザに応じて異なっており、相互利用が容易な形式で保存する必要がある。本実施形態では、所定ルールに従い各ユーザが文書又は当該文書に関連するファイルを生成する。生成された情報を情報処理装置 1 により一元管理することで、スムーズなデータ共有を実現する。

[0017] 情報処理装置 1 は、コンピュータであり、図 1 に示すように、制御部 1 1、記憶部 1 2、及び通信部 1 3 を備える。情報処理装置 1 は複数のコンピュータからなるマルチコンピュータであってもよく、ソフトウェアによって仮想的に構築された仮想マシンであってもよい。

[0018] 制御部 1 1 は、一又は複数の CPU (Central Processing Unit)、MPU (Micro-Processing Unit)、GPU (Graphics Processing Unit) 等の演算処理装置を備える。制御部 1 1 は、内蔵する ROM (Read Only Memory) 又は RAM (Random Access Memory) 等のメモリ、クロック、カウンタ等を用

い、各構成部を制御して処理を実行する。なお、情報処理装置 1 の機能は、ソフトウェア的に実現してもよいし、一部又は全部を、例えば A S I C (Application Specific Integrated Circuit)、F P G A (Field Programmable Gate Array) 等のハードウェアで実現してもよい。

[0019] 記憶部 1 2 は、例えばハードディスク、フラッシュメモリ、S S D (Solid State Drive) 等の不揮発性メモリを備える。記憶部 1 2 は、情報処理装置 1 に接続された外部記憶装置であってもよい。記憶部 1 2 は、制御部 1 1 が参照する各種コンピュータプログラム及びデータを記憶する。記憶部 1 2 は、文書の提供に関する処理をコンピュータに実行させるためのプログラム 1 P と、当該プログラム 1 P の実行に必要なデータとしての文書 D B (Data Base : データベース) 1 2 1 とを記憶している。文書 D B 1 2 1 は、各ユーザから受け付けた文書に関する情報を格納するデータベースである。

[0020] プログラム 1 P を含むコンピュータプログラム (コンピュータプログラム製品) は、当該コンピュータプログラムを読み取り可能に記録した非一時的な記録媒体 1 A により提供されてもよい。記憶部 1 2 は、不図示の読出装置によって記録媒体 1 A から読み出されたコンピュータプログラムを記憶する。記録媒体 1 A は、例えば磁気ディスク、光ディスク、半導体メモリ等である。また、通信ネットワークに接続されている外部サーバからコンピュータプログラムをダウンロードし、記憶部 1 2 に記憶させてもよい。プログラム 1 P は、単一のコンピュータプログラムでも複数のコンピュータプログラムにより構成されるものでもよく、また、単一のコンピュータ上で実行されても通信ネットワークによって相互接続された複数のコンピュータ上で実行されてもよい。

[0021] 通信部 1 3 は、ネットワーク N を介した通信を実現する通信デバイスを備える。制御部 1 1 は、通信部 1 3 を介して各端末装置 2 との間でデータを送受信する。

[0022] 情報処理装置 1 の構成は上述の例に限定されず、例えばユーザの操作を受け付けるための操作部、画像を表示する表示部等を備えてもよい。

- [0023] 端末装置 2 は、コンピュータであり、制御部 2 1、記憶部 2 2、通信部 2 3、表示部 2 4、及び操作部 2 5 を備える。
- [0024] 制御部 2 1 は、一又は複数の CPU、MPU、GPU 等の演算処理装置を備える。制御部 1 1 は、内蔵する ROM 又は RAM 等のメモリ、クロック、カウンタ等を用い、各構成部を制御して処理を実行する。
- [0025] 記憶部 2 2 は、例えばハードディスク、フラッシュメモリ、SSD 等の不揮発性メモリを備える。記憶部 2 2 は、制御部 2 1 が参照する各種コンピュータプログラム及びデータを記憶する。記憶部 2 2 は、文書の検索に関する処理をコンピュータに実行させるためのプログラム 2 P を記憶している。
- [0026] 通信部 2 3 は、ネットワーク N を介した通信を実現する通信デバイスを備える。制御部 2 1 は、通信部 2 3 を介して情報処理装置 1 との間でデータを送受信する。
- [0027] 表示部 2 4 は、例えば液晶ディスプレイ、有機 EL (Electro Luminescence) ディスプレイ等のディスプレイ装置を備える。表示部 2 4 は、制御部 2 1 からの指示に従い、情報処理装置 1 から受信した文書等を表示する。
- [0028] 操作部 2 5 は、ユーザの操作を受け付けるインタフェースである。操作部 2 5 は、例えばキーボード、マウス、ディスプレイ内蔵のタッチパネルデバイス、スピーカ及びマイクロフォン等を備える。操作部 2 5 は、ユーザからの操作入力を受け付け、操作内容に応じた制御信号を制御部 2 1 へ送出する。
- [0029] 図 2 は、文書 DB 1 2 1 に記憶される情報の内容例を示す図である。文書 DB 1 2 1 には、例えばデータ ID をキーに、文書情報、まとめファイル情報、及び実験ファイル情報等の情報を紐付けたレコードが格納されている。
- [0030] 文書情報は、ユーザにより作成された文書に関する情報を含む。文書情報には、例えば、文書を識別する文書 ID、文書ファイル（文書データ）、及び文書ファイルが更新された日付を表す更新年月日等が含まれる。文書ファイル列には、実体ファイルが格納される。
- [0031] 文書ファイルには、表紙と、本文とが含まれる。表紙は、予め設定される

形式に従い、所定項目が記載されたページである。本文は、例えばテキスト、画像、音声等のデータを含み、文書の作成者の構文や言語等に基づく自由形式のデータである。すなわち、本実施形態における文書ファイルの本文は、非構造化データである。非構造化データとは、計算機で統計処理が行えるような意味づけが含まれてないデータを意味する。なお、文書ファイルは、表紙及び本文からなる一連の文書であるものに限らず、表紙ファイル及び本文ファイルにより構成されてもよい。

[0032] まとめファイル情報は、まとめファイルに関する情報を含む。まとめファイルとは、文書に対応付けて作成されるものであり、文書に関する所定の情報を整理してまとめ上げたファイルである。まとめファイル情報には、例えば、まとめファイルを識別するまとめファイルID、まとめファイル（まとめデータ）、及びリンク情報等が含まれる。まとめファイル列には、実体ファイルが格納される。リンク情報は、まとめファイルへアクセスするための当該まとめファイルへのリンク情報（パス情報）を含む。なお、1つの文書ファイルに紐付けられるまとめファイルの数は1又は3以上であってもよく、ゼロであってもよい。

[0033] 実験ファイル情報は、実験ファイルに関する情報を含む。実験ファイルは、文書に関連する実験の実験データを表す。実験ファイル情報には、例えば、実験ファイルを識別する実験ファイルID、実験ファイル（実験データ）、及びリンク情報等が含まれる。実験ファイル列には、実体ファイルが格納される。実験ファイルは、例えば、テキスト、グラフ、画像等のデータを含み、非構造化データである。リンク情報は、実験ファイルへアクセスするための当該実験ファイルへのリンク情報（パス情報）を含む。なお、1つの文書ファイルに紐付けられる実験ファイルの数は1又は3以上であってもよく、ゼロであってもよい。

[0034] 文書DB121は、例えば階層化されたファイル構造を有し、当該階層構造に合わせて文書セットを記憶する。一例として、文書DB121は、第1階層から第3階層で構成される。第1階層は、全社フォルダで構成される。

第2階層は、会社の研究所それぞれに対応付けられた複数のフォルダで構成される。研究所フォルダの下階層である第3階層は、研究所の部署としてのグループそれぞれに対応付けられた一又は複数のフォルダと、研究所におけるプロジェクトそれぞれに対応付けられた一又は複数のフォルダとで構成される。各グループフォルダに、グループに対応する文書セットが格納される。また各プロジェクトフォルダに、プロジェクトに対応する文書セットが格納される。なお、上述の階層構造は単なる例示であり、文書DB121の文書の管理体制は組織の構成や文書の管理目的に応じて適宜設計されてよい。例えば、プロジェクトフォルダは第2階層に含まれてもよい。文書DB121は階層型でなくてもよい。

[0035] 文書セットを格納するフォルダは、例えば文書の種類や作成者の所属部署に基づき情報処理装置1が自動的に判定する。なお文書セットの格納先は、文書セットを端末装置2aから取得する際に第1ユーザからの指定を受け付けることにより決定されてもよい。

[0036] 第1ユーザは、作成した文書を本システムに提供する際、端末装置2aを用いて、予め設定された所定形式に従い必要項目を入力し、表紙を作成する。また、必要に応じて、予め設定された所定形式に従い実験結果又は研究結果等を入力して、まとめファイルを作成する。第1ユーザは、文書の本文に表紙を添えた文書ファイルに、生成したまとめファイル及び予め作成（保存）済みの実験ファイルに対応付けた文書セットを情報処理装置1へ送る。

[0037] 情報処理装置1は、例えば、端末装置2aから文書セットのアップロードを受け付けてもよく、端末装置2aから送信された電子メールやチャットメッセージ等のメッセージを受信することにより、当該メッセージに添付された文書セットを取得してもよい。情報処理装置1は、取得した文書セットを文書DB121に記憶する。文書DB121の内容は、随時更新される。

[0038] 図3は、表紙の内容例を示す図である。図3に示すように、表紙に記載される事項には、例えば、書誌情報、まとめファイル保存場所、実験ファイル保存場所等が含まれる。書誌情報は、文書を特定するための情報であり、例

例えば、文書のタイトル、作成者氏名、作成者の所属先、作成者の所属チーム、作成者の連絡先、文書の作成年月日等を含む。まとめファイル保存場所は、文書に関連するまとめファイルの保存場所（リンク情報）を表す。まとめファイル保存場所におけるファイル名及びリンク情報の記載有無により、文書に関連するまとめファイルの存否を把握可能である。実験ファイル保存場所は、実験ファイルの保存場所（リンク情報）を表す。実験ファイル保存場所におけるファイル名及びリンク情報の記載有無により、文書に関連する実験ファイルの存否を把握可能である。

[0039] 図4は、まとめファイルの内容例を示す図である。まとめファイルは、例えば文書に関連する実験結果や研究結果等のデータを、予め定義された構造に従って記述、配置したデータ集合である。すなわち、まとめファイルは構造化データである。構造化データとは、一定の構造（規則）に従って記述され、計算機で統計処理が行えるような意味づけが予め含まれているデータを意味する。

[0040] 図4に示す例にて、まとめファイルは、「列」及び「行」を有する表形式であり、「列」等に意味を持たせて構造化したものである。まとめファイルは、項目として、サンプル（物質）を識別するためのサンプルID、注目物性及び物質情報を含む。注目物性は、まとめファイルの複数のデータのうちの注目すべき物性であり、例えば物性評価の対象となった物性である。物質情報は、物質に関する情報であって、注目物性と関係性を有し得る情報を含む。物質情報は、例えば物質の物性を表す物性値、構造を表す構造データ、物質におけるプロセスに関するプロセス値等を含んでもよい。

[0041] 本実施形態のまとめファイルには、注目物性としての実験値1と、物質情報としての実験値2、計算値1、計算値2、プロセス値1、及びプロセス値2とが含まれている。なお、注目物性及び物質情報に含まれるデータ数及びデータ種類は単なる例示である。注目物性には、複数種類の物性が含まれていてもよい。

[0042] 実験値は、物性値に対応し、実験により得られた値である。計算値は、物

性値又は構造データに対応し、例えば実験値や他の計算値等に基づく計算により得られた値である。プロセス値には、例えば物質に関する各種プロセスのプロセス条件が含まれる。プロセス値としては、例えば、物質を合成する合成工程における合成条件、物質を作製（調製）する作製工程における作製条件等が挙げられる。

[0043] 計算値は、物性を表すシミュレーション値であってもよい。シミュレーション値は、所定のアルゴリズムにより求められた物質の物性値である。シミュレーション値は、例えば、量子化学計算、分子動力学計算、第一原理計算等の理論計算により得られる。計算値は、物質の構造を表す分子記述子であってもよい。分子記述子は、物質の持つ構造的特徴や物理化学的特性等を計算機で扱いやすくするために数値化したものである。分子記述子は、物質の構造式から計算可能であり、公知のソフトウェア、例えばRDKit、mordred、MOE、alvaDesc、PaDEL-Descriptor、Codessa等を用いて求めることができる。

[0044] まとめファイルの内容は上述の例に限らず、技術文書の内容に応じて適宜の項目を設定することができる。

[0045] 図5は、検索システム100が実行する文書の取得処理手順の一例を示すフローチャートである。以下のフローチャートにおける処理は、情報処理装置1の記憶部12に記憶するプログラム1Pに従って制御部11によって実行される。情報処理装置1は、例えば端末装置2aから文書セットが送信されると、以下の処理を実行する。

[0046] 情報処理装置1の制御部11は、端末装置2aから文書ファイル、まとめファイル及び実験ファイルに対応付けた文書セットを取得する（ステップS11）。まとめファイル及び実験ファイルの一方又は両方は文書ファイルに対応付けられていなくてもよい。

[0047] 制御部11は、取得した文書ファイル、まとめファイル及び実験ファイルに対応付けて文書DB121に記憶する（ステップS12）。ステップS12において制御部11は、例えば、取得した文書ファイルの種類や作成者の

所属部署に基づき文書セットの格納先を決定し、決定した格納先のフォルダへ文書セットを格納する。制御部 11 は、決定した格納先に基づきまとめファイルのリンク先及び実験ファイルのリンク先を特定し、特定したリンク先を文書 DB 121 に記憶する。制御部 11 はまた、特定したリンク先を、表紙のまとめファイル保存場所及び実験ファイル保存場所に追加する。なお、表紙のファイル保存場所及び実験ファイル保存場所は、第 1 ユーザ側で予め入力されていてもよい。この場合、制御部 11 は、入力された保存場所に従い、当該保存場所にまとめファイル及び実験ファイルを格納するとともに、各保存場所を文書 DB 121 に記憶する。

[0048] まとめファイル及び実験ファイルの一方又は両方は、文書 DB 121 に保存されるものに限らず、第 1 ユーザの端末装置 2a に記憶されていてもよい。この場合、情報処理装置 1 は、文書ファイルと、まとめファイル及び実験ファイルにアクセスするためのリンク先とを対応付けて端末装置 2a から取得してもよい。

[0049] 各端末装置 2a から文書セットが送信される度、上述の処理が繰り返し実行されることにより、社内における複数のユーザにより作成された多種多様な文書が文書 DB 121 に蓄積される。

[0050] 図 6 は、検索システム 100 が実行する検索処理手順の一例を示すフローチャートである。以下の各フローチャートにおける処理は、情報処理装置 1 の記憶部 12 に記憶するプログラム 1P に従って制御部 11 によって実行されるとともに、端末装置 2b の記憶部 12 に記憶するプログラム 2P に従って制御部 21 によって実行される。端末装置 2b の制御部 21 は、例えば、プログラム 2P が第 2 ユーザによって選択されて起動され、メニューから「文書検索」が選択されると、以下の処理を実行する。

[0051] 端末装置 2b の制御部 21 は、第 2 ユーザの操作部 25 の操作による文書の検索条件及び検索要求を受け付ける（ステップ S21）。ステップ S21 では、制御部 21 は、検索条件を受け付ける受付画面を表示部 24 に表示させ、受付画面を利用して文書の検索に対する検索条件及び検索要求を受け付

ける。

[0052] 図7は、端末装置2bの表示部24に表示される受付画面31の一例を示す模式図である。端末装置2bの制御部21は、例えば、情報処理装置1が提供する文書検索サービスに対する第2ユーザからのアカウント情報を用いたログインを受け付けた場合、情報処理装置1から送信される画面情報に基づき、図7に示す受付画面31を表示させる。

[0053] 受付画面31には、例えば、文書に関する検索キーワードの入力を受け付けるキーワード欄310と、絞り込み条件の入力を受け付ける絞り込み条件欄311と、検索要求の指示の入力を受け付ける検索ボタン312とが含まれる。

[0054] キーワード欄310は、文字入力欄を有する。第2ユーザは、操作部25を用いてキーワード欄310に検索キーワードを入力することにより、所望の検索キーワードを含む文書を検索することが可能である。

[0055] 絞り込み条件欄311は、例えば複数の絞り込み条件項目に対する絞り込み条件を受け付け可能に構成される。図7では、絞り込み条件欄311は、更新日の絞り込み条件の指定を受け付ける第1領域311aと、検索フォルダ範囲の絞り込み条件の指定を受け付ける第2領域311bとを備える。第1領域311a及び第2領域311bには、各絞り込み条件項目それぞれに対応付けられた複数のチェックボックス313が設けられている。第2ユーザは、操作部25を用いて絞り込み条件欄311における一又は複数のチェックボックス313にチェックを入力することにより、所望の絞り込み条件項目に該当する文書を検索することが可能である。

[0056] 更新日の絞り込み条件項目としては、例えば、「5日以内」、「2週間以内」、「1か月以内」等が挙げられる。検索フォルダ範囲の絞り込み条件項目としては、例えば、「全社」、「〇〇研究所」、「〇〇グループ」、「〇〇プロジェクト」等が挙げられる。検索範囲は、上述した文書DB121における各階層のフォルダ名に対応する。

[0057] 絞り込み条件欄311の構成は、絞り込み条件項目を受け付け可能であれ

ば特に限定されない。絞り込み条件欄 3 1 1 は、予め設定される絞り込み条件項目候補の中から選択を受け付けるものに限らず、第 2 ユーザからの入力に応じて任意の絞り込み条件項目、例えば任意の更新期間、任意のフォルダ範囲の指定を受け付け可能に構成されてもよい。

[0058] なお、文書 DB 1 2 1 は、例えば階層毎及び各階層に作成されたフォルダ毎のように、所定基準でアクセス権が定義されていてもよい。この場合、情報処理装置 1 は、端末装置 2 b への受付画面 3 1 の提供に際し、ログインされたアカウント情報に基づき、文書検索を行う第 2 ユーザのアクセス権を特定する。情報処理装置 1 は、第 2 ユーザに許容されたアクセス権の範囲内におけるフォルダのみを、検索フォルダ範囲の絞り込み条件項目として領域に表示させることができる。

[0059] 絞り込み条件欄 3 1 1 は、図 7 に示すように、絞り込み項目を選択する第 3 領域 3 1 1 c を備え、ユーザの選択に応じて、絞り込み項目の選択、変更、追加等を受け付け可能に構成されてもよい。第 3 領域 3 1 1 c は、例えばプルダウン式のメニューを有する。第 2 ユーザは、プルダウン式のメニューを開くことによって、予め用意されている複数の絞り込み条件の中から一つの絞り込み条件を選択することが可能である。いずれかの絞り込み条件が選択された場合、第 3 領域 3 1 1 c は、第 1 領域 3 1 1 a 及び第 2 領域 3 1 1 b と同様に、絞り込み条件項目それぞれに対応付けられた複数のチェックボックス 3 1 3 を表示するよう構成されてもよい。

[0060] キーワード欄 3 1 0 及び絞り込み条件欄 3 1 1 に検索キーワード及びチェックが入力された状態で、検索ボタン 3 1 2 を第 2 ユーザが指定することにより、操作部 2 5 に検索条件及び文書の検索の要求が入力される。制御部 2 1 は、検索条件及び文書の検索の要求を受け付ける。なお第 2 ユーザは、検索キーワード及び絞り込み条件のいずれか一方のみを検索条件として指定してもよい。

[0061] 図 6 に戻り説明を続ける。制御部 2 1 は、受け付けた検索条件及び文書の検索の要求を情報処理装置 1 に送信する（ステップ S 2 2）。

- [0062] 情報処理装置 1 の制御部 1 1 は、端末装置 2 b から検索条件及び文書の検索の要求を受信する（ステップ S 2 3）。検索条件には、例えば検索キーワードと、一又は複数の絞り込み条件とが含まれる。
- [0063] 制御部 1 1 は、文書 DB 1 2 1 に記憶する情報に基づき、文書 DB 1 2 1 に保存される文書のうち、受信した検索条件を満たす文書を特定する（ステップ S 2 4）。例えば、検索キーワードを取得した場合、制御部 1 1 は、文書ファイルに取得した検索キーワードが含まれる文書を特定する。更新日の絞り込み条件を取得した場合、制御部 1 1 は、文書情報における更新年月日が、取得した更新日の絞り込み条件の期間に含まれる文書を特定する。検索フォルダ範囲の絞り込み条件を取得した場合、制御部 1 1 は、取得した検索フォルダ範囲の範囲内に限定して文書を検索する。制御部 1 1 は、文書 DB 1 2 1 の検索を実行し、全ての検索条件を満たす文書を抽出する。
- [0064] 制御部 1 1 は、特定した検索条件を満たす各文書の概要情報を含む概要画面を生成し（ステップ S 2 5）、生成した概要画面をログインした第 2 ユーザの端末装置 2 b に送信する（ステップ S 2 6）。概要情報は、文書の概要を表す情報を含む。
- [0065] 端末装置 2 b の制御部 2 1 は、概要画面を情報処理装置 1 から受信し（ステップ S 2 7）、受信した概要画面を表示部 2 4 に表示する（ステップ S 2 8）。
- [0066] 図 8 は、端末装置 2 b の表示部 2 4 に表示される概要画面 3 2 の一例を示す模式図である。概要画面 3 2 には、例えば、文書に関する検索条件を表示する検索条件欄 3 2 0 と、文書の概要情報を表示する概要情報欄 3 2 1 とが含まれる。
- [0067] 検索条件欄 3 2 0 は、概要情報欄 3 2 1 に表示される文書に対応する検索条件を表示する。情報処理装置 1 の制御部 1 1 は、文書の検索に用いた検索キーワード及び絞り込み条件等の検索条件を検索条件欄 3 2 0 に表示させる。検索条件欄 3 2 0 は、受付画面 3 1 としての機能を備え、当該検索条件欄 3 2 0 上で検索条件の変更、追加等を受け付け可能に構成されてもよい。

- [0068] 概要情報欄 3 2 1 は、検索条件欄 3 2 0 に示される検索条件を満たす各文書の概要情報を一覧で表示する。概要情報とは、文書の概要を表す情報である。図 8 に示す例にて、概要情報欄 3 2 1 は、文書の表紙のサムネイルを表示する第 4 領域 3 2 1 a と、文書の一部を抜粋した抜粋情報を表示する第 5 領域 3 2 1 b とを含み、概要情報としてのサムネイル及び抜粋情報を表示する。
- [0069] 情報処理装置 1 の制御部 1 1 は、検索条件を満たす文書を特定した場合、特定された各文書について、文書 DB 1 2 1 から表紙を読み出し、読み出した表紙のサムネイルを生成し、生成したサムネイルを第 4 領域 3 2 1 a に表示させる。なお文書のサムネイルは表紙以外の部分のサムネイルであってもよい。制御部 1 1 は、例えば検索キーワードを含むページのサムネイル、本文の 1 ページ目のサムネイル等を生成してもよい。
- [0070] 制御部 1 1 はまた、文書 DB 1 2 1 から文書ファイルを読み出し、読み出した文書ファイルのうち、例えば検索キーワードを含む所定長のテキストデータを抜粋し、抜粋したテキストデータの文字を抜粋情報として第 5 領域 3 2 1 b に表示させる。制御部 1 1 は、検索キーワードに対応する文字部分に対し、ハイライト表示等の所定の表示処理を施してもよい。なお、検索条件として検索キーワードを取得していない場合には、初期設定される特定部分（例えば本文の最初の一行）が抜粋されてもよい。
- [0071] 概要情報欄 3 2 1 に一覧表示される各概要情報は、選択可能に構成されている。概要情報欄 3 2 1 は、当該概要情報欄 3 2 1 に表示される文書に係る概要情報のうちのいずれかの文書に係る概要情報の選択を受け付ける文書選択欄としても機能する。第 2 ユーザは、操作部 2 5 を用いていずれかの文書に対応する第 4 領域 3 2 1 a 又は第 5 領域 3 2 1 b を選択（指定）することにより、さらなる詳細情報の取得を要求する文書を選択することが可能である。
- [0072] 詳細情報とは、文書の詳細を表す情報であり、概要情報よりも多くの情報量を含む。本明細書において、詳細情報とは、例えば表紙及び本文を含む文

書ファイルのように文書自体の詳細に加え、まとめファイル及び実験ファイルのように文書に紐付けられたファイルの詳細も、その意味に含む。

[0073] 図8に示す如く概要画面32が表示された状態で、いずれかの第4領域321a又は第5領域321bを第2ユーザが指定することにより、操作部25に特定の文書の選択及び当該特定の文書に対する詳細情報の要求が入力される。制御部21は、特定の文書の選択及び詳細情報の要求を受け付ける。詳細情報の要求に対応する文書の特定は、例えば指定された第4領域321a又は第5領域321bに基づき制御部21が自動的に判定する。制御部21は、例えば操作部25に入力された概要画面32上の位置座標に基づき、指定された第4領域321a又は第5領域321bを特定し、特定した第4領域321a又は第5領域321bに対応するサムネイル又は抜粋情報を認識する。制御部21は、認識したサムネイル又は抜粋情報のメタデータ等に基づき、詳細情報の要求に対応する文書の文書IDを特定する。

[0074] 図6に戻り説明を続ける。制御部21は、特定の文書の選択及び当該特定の文書に対する詳細情報の要求を受け付ける（ステップS29）。制御部21は、受け付けた特定の文書の選択及び詳細情報の要求を情報処理装置1に送信する（ステップS30）。

[0075] 情報処理装置1の制御部11は、特定の文書の選択及び詳細情報の要求を受信する（ステップS31）。なお、詳細情報の要求に対応する文書の特定は、情報処理装置1側で行ってもよい。この場合、情報処理装置1は、端末装置2bの操作部25を通じて受け付けた文書の選択に係る位置座標を示す情報を端末装置2bから受信し、受信した位置座標に基づき詳細情報の要求に対応する文書を特定してもよい。

[0076] 制御部11は、文書DB121に記憶する情報に基づき、詳細情報が要求された特定の文書の詳細情報を取得する（ステップS32）。制御部11は、例えば、文書の詳細として特定の文書の文書IDに対応する文書ファイルを文書DB121から抽出する。制御部11は、取得した文書ファイルを含む詳細画面を生成する（ステップS33）。制御部11は、生成した詳細画

面をログインした第2ユーザの端末装置2bに送信する（ステップS34）。

[0077] 端末装置2bの制御部21は、詳細画面を情報処理装置1から受信し（ステップS35）、受信した詳細画面を表示部24に表示する（ステップS36）。詳細画面には、文書ファイルが表示される。詳細画面は、例えばスクロール又はページめくり機能等により、表紙及び本文の各ページを確認可能に表示する。表紙には、上述の通り文書に関連するまとめファイル又は実験ファイルが存在する場合における各ファイルへのリンク情報が表示されている。

[0078] 制御部21は、まとめファイル又は実験ファイルを取得するか否かを判定する（ステップS37）。第2ユーザの操作による表紙のリンク情報の選択を受け付けていないことにより、まとめファイル又は実験ファイルを取得しないと判定した場合（S37：NO）、制御部21は処理を終了する。

[0079] 第2ユーザの操作による表紙のリンク情報の選択を受け付けたことにより、まとめファイル又は実験ファイルを取得すると判定した場合（S37：YES）、制御部21は、情報処理装置1にまとめファイル又は実験ファイルを要求する（ステップS38）。詳細には、制御部21は、情報処理装置1に選択されたリンク情報のリンク先に格納されるまとめファイル又は実験ファイルを要求する。

[0080] 情報処理装置1は、まとめファイル又は実験ファイルの要求を端末装置2bから受け付ける（ステップS39）。情報処理装置1は、選択されたリンク情報に示される保存場所に記憶するまとめファイル又は実験ファイルを文書DB121から読み出し、読み出したまとめファイル又は実験ファイルを端末装置2bに送信する（ステップS40）。

[0081] 端末装置2bの制御部21は、まとめファイル又は実験ファイルを情報処理装置1から受信し（ステップS41）、受信したまとめファイル又は実験ファイルを表示部24に表示する（ステップS42）。なお、情報処理装置1以外の他装置（例えば他の端末装置2）にまとめファイル又は実験ファイ

ルが格納されている場合には、制御部 21 は、リンク情報に従い、他装置にまとめファイル又は実験ファイルを要求してもよい。制御部 21 は情報処理装置 1 を介して他装置にまとめファイル又は実験ファイルを要求してもよい。制御部 21 は処理を終了する。制御部 21 は、処理をステップ S 21、S 28、S 36 等へ戻し、操作受付を続行してもよい。

[0082] 本実施形態によれば、複数のユーザ間で生成された技術文書を情報処理装置 1 により一元管理することで、各ユーザが検索画面を用いて所望の文書を容易に検索、取得することができ、データ共有の利便性が向上される。共有される文書には、所定フォーマットの表紙が添えられているため、全てのユーザが一見して容易且つ明確にその文書の内容を把握することができる。

[0083] 文書を提供するユーザは、用意されたフォーマットを利用することで、表紙の生成や文書の登録を容易に行うことができる。

[0084] 文書には、文書に関連するまとめファイルや実験ファイルが紐付けられているため、文書の全内容に加えて、より多くの情報を効率的に取得することができる。まとめファイルや実験ファイルといった添付資料自体の内容は直接の検索対象とせず、検索対象を文書のみにするすることで、検索負荷を低減し、処理性能が向上される。

[0085] 文書に関連する各種データに基づき構造化されたまとめファイルが生成されることで、各ユーザが個々に取得したデータの相互運用性を高めることができる。構造化データとしてデータを保存することで、効率的なデータ把握や分析が可能であり、データ連携を進める上で発生するデータの認識齟齬を低減することができ、データの利活用性が高まる。

[0086] 文書検索に際し、検索キーワードや絞り込み条件を設定することができるため、所望の文書を効率的に検索することができる。概要画面 32 には、文書のサムネイルや一部抜粋が一覧表示されるため、検索結果を一見して把握することができる。また、表紙に添付資料へのリンク情報を埋め込むことで、ユーザは容易且つ確実に添付資料の有無を把握できるとともに、必要に応じて添付資料を容易に入手することができ、使用性が向上される。

[0087] (第2実施形態)

第2実施形態では、まとめファイルを用いて学習モデルを構築する構成を説明する。以下の実施形態では主に第1実施形態との相違点を説明し、第1実施形態と共通する構成については同一の符号を付してその詳細な説明を省略する。

[0088] 図9は、第2実施形態の検索システム100の構成例を示すブロック図である。第2実施形態における情報処理装置1の記憶部12には、学習モデル122が記憶されている。学習モデル122は、機械学習により生成された学習モデルである。学習モデル122は、人工知能ソフトウェアの一部を構成するプログラムモジュールとしての利用が想定される。

[0089] 図10は、学習モデル122の概要を示す説明図である。図10では、学習モデル122として、物質に関する物質情報を入力として、当該物質の物性を示す情報を出力するモデルを例示する。

[0090] 学習モデル122は、例えばニューラルネットワークである。学習モデル122は、物質情報が入力される入力層と、物性値を出力する出力層と、特徴量を抽出する中間層（隠れ層）とを備える。中間層は、畳み込み層、プーリング層及び全結合層等を含んでもよい。中間層は、入力データの特徴量を抽出する複数のノードを有し、各種パラメータを用いて抽出された特徴量を出力層に受け渡す。入力層に物質情報が入力された場合、学習済みパラメータによって中間層で演算が行なわれ、出力層から、物性を示す出力情報が出力される。

[0091] 学習モデル122に入力される物質情報は、例えば物質の構造を表す構造データ、物性を表す物性値、プロセスに関するプロセス値等を含む。構造データ、物性値及びプロセス値にはそれぞれ、複数種類のデータが含まれてもよい。学習モデル122には、構造データ、物性値及びプロセス値の少なくとも1つが入力されてもよい。

[0092] 学習モデル122の出力層は1つのノードを有し、当該ノードから所定の物性を示す連続値を出力する。なお、予測すべき物性に関する情報を得るこ

とができれば、出力層の構成は特に限定されない。学習モデル 1 2 2 は、複数種類の物性を予測可能に構成されてもよい。

[0093] 学習モデル 1 2 2 は、物質の物質情報に対し、既知の物性値を示すデータがラベリングされた訓練データを用意し、当該訓練データを用いて未学習のニューラルネットワークを機械学習させることにより生成することができる。

[0094] 本実施形態では、上記訓練データとして、まとめファイルのデータを利用する。まとめファイルには、図 4 に示したように、複数の物質についての構造データ、物性値又はプロセス値に相当する実験値、計算値及びプロセス値が、構造化データとして記憶されている。従って、例えば、まとめファイルにおける物質情報列のデータを入力情報とし、注目物性列のデータを出力情報とすることで、物質情報に対し、正解値となる物性値が付与された複数のデータセットからなる訓練データを容易に得ることができる。

[0095] 図 1 1 は、第 2 実施形態の検索システム 1 0 0 が実行する処理手順の一例を示すフローチャートである。端末装置 2 b の制御部 2 1 は、例えば、プログラム 2 P が第 2 ユーザによって選択されて起動され、メニューから「機械学習」が選択されると、以下の処理を実行する。

[0096] 端末装置 2 b の制御部 2 1 は、第 2 ユーザの操作を受け付けることによるいずれかのまとめファイルへのリンク情報を取得し、取得したリンク先に格納されるまとめファイルを取得する（ステップ S 5 1）。制御部 2 1 は、取得したまとめファイルを表示部 2 4 に表示する（ステップ S 5 2）。第 2 ユーザは、まとめファイルをそのまま訓練データとしてもよく、あるいは操作部 2 5 を用いてまとめファイルのデータのうちの一部を削除又は他のデータを追加等することにより、所望のデータセットからなる訓練データを生成することが可能である。

[0097] 制御部 2 1 は、まとめファイルを用いて生成された訓練データを取得する（ステップ S 5 3）。制御部 2 1 は、表示したまとめファイルに対し、訓練データにおける入力情報となるデータ列と、出力情報となるデータ列との選

択を第2ユーザから受け付けることにより、入力情報と出力情報とを対応付けた訓練データとして取得してもよい。制御部21は、取得した訓練データを情報処理装置1に送信する（ステップS54）。制御部21は、所定の登録画面に訓練データをアップロードしてもよい。

[0098] 情報処理装置1の制御部11は、端末装置2bから訓練データを受信する（ステップS55）。制御部11は、受信した訓練データを用いて、物質の物質情報を入力した場合に、当該物質の所定物性の物性値を出力する学習モデル122を生成する（ステップS56）。

[0099] 具体的には、制御部11は、訓練データに含まれる物質情報を学習モデル122の入力層に入力し、中間層での演算処理を経て、出力層から出力される物性値を取得する。制御部11は、出力層から出力された物性値と、訓練データに含まれる物性値とを比較し、出力層から出力される物性値が正解値に近づくように、ニューロン間の重み（結合係数）等のパラメータを最適化する。パラメータの最適化の方法は特に限定されないが、例えば制御部11は誤差逆伝播法を用いて各種パラメータの最適化を行う。学習が開始される前の段階では、学習モデル122を記述する定義情報には、初期設定値が与えられているものとする。誤差、学習回数等が所定基準を満たすことによって学習が完了すると、最適化されたパラメータが得られる。

[0100] 学習が完了すると、制御部11は、学習済みの学習モデル122として、学習済みの学習モデル122に関する定義情報を記憶部12に記憶させ（ステップS57）、一連の処理を終了する。上述の処理により、物質の物質情報に対し所定物性の物性値を適切に予測可能に学習された学習モデル122が構築される。

[0101] 情報処理装置1は、構築した学習モデル122を用いて、各物質の物性予測処理を実行することができる。情報処理装置1は、例えば第2ユーザの端末装置2bから、物性予測対象となる物質の物質情報を受信する。情報処理装置1は、いずれかのまとめファイルにおける所定列の指定を受け付けることにより、物性予測対象となる物質の物質情報を取得してもよい。情報処理

装置 1 は、取得した物質の物質情報を学習モデル 1 2 2 に入力して、学習モデル 1 2 2 から出力される物性値を取得する。情報処理装置 1 は、得られた物性値を端末装置 2 b へ送信する。なお、生成された学習モデル 1 2 2 は、端末装置 2 b にデプロイされ、端末装置 2 b 側で予測処理が実行されてもよい。

[0102] 上述の処理において、プログラム 1 P 及びプログラム 2 P の少なくとも一方は、例えば訓練データを生成するための E T L ツール及び機械学習を実行するための機械学習ツールを含む、又はこれらとアプリケーション間連携するものであってもよい。E T L ツールとは、データベースに蓄積されたデータを抽出 (Extract) し、データウェアハウスをはじめとするソフトウェアで利用しやすい形に変換 (Transform) し、ソフトウェアに書き出す (Load) という一連の処理を実行するツールである。例えば第 2 ユーザは、E T L ツールを利用してまとめファイルのデータを加工することにより、訓練データを生成してもよい。

[0103] 上述の処理において、情報処理装置 1 は、訓練データの生成から学習モデル 1 2 2 の構築までを一連の処理として自動で実行してもよい。情報処理装置 1 は、例えば、第 2 ユーザの操作を受け付けることによるいずれかのまとめファイルを指定した学習モデル 1 2 2 の生成要求を受け付ける。情報処理装置 1 は、まとめファイルのデータを解析して、入力情報となるデータ列及び出力情報となるデータ列を認識することにより、訓練データを生成する。情報処理装置 1 は、まとめファイルにおける物質情報列を入力情報とし、注目物性列を出力情報と特定してもよい。情報処理装置 1 は、生成した訓練データを用いて、学習モデル 1 2 2 を構築する。

[0104] 訓練データには、まとめファイルにおける一部のデータ列のみが選択的に使用されてもよい。例えば、注目物性と各物質情報項目との相関関係に基づいて、注目物性と相関が比較的強い物質情報項目のみを説明変数として採用してもよい。又は、物質情報項目同士の相関関係に基づいて、互いに相関が強い複数の物質情報項目のうちの所定数の物質情報項目のみを説明変数とし

て採用し、残りの物質情報項目の物質情報を削除することで、データ量を削減してもよい。相関関係は、例えば過去の知見から特定されてもよく、所定の相関係数を求める算式を用いて、まとめファイルのデータから求めてもよい。情報処理装置 1 は、まとめファイルに基づき相関関係の度合い（例えば相関係数の値）を求め、求めた相関関係の度合いと、訓練データ生成条件とに基づき訓練データを生成する。情報処理装置 1 は、求めた相関関係の度合いを端末装置 2 b に提供してもよい。第 2 ユーザは、提供された相関関係の度合いを確認して、訓練データに使用するデータを選択することができる。

[0105] 本実施形態によれば、まとめファイルを利用して容易に学習モデル 1 2 2 を生成することができる。共有データから検索された文書に基づき、まとめファイルを効率的に取得することができるため、ユーザの利便性が向上する。まとめファイルは予め構造化データとして生成されているため、煩雑な処理を行うことなく学習モデル 1 2 2 の学習処理に使用することができる。まとめファイルを利用することで、複数のユーザにより得られた複数の実験結果を容易に連携させることができるため、訓練データの品質が向上し、精度のよい学習モデル 1 2 2 の生成につながる。

[0106] 今回開示した実施の形態は、全ての点で例示であって、制限的なものではないと考えられるべきである。各実施例にて記載されている技術的特徴は互いに組み合わせることができ、本発明の範囲は、請求の範囲内の全ての変更及び請求の範囲と均等の範囲が含まれることが意図される。

各実施形態に示すシーケンスは限定されるものではなく、矛盾の無い範囲で、各処理手順はその順序を変更して実行されてもよく、また並行して複数の処理が実行されてもよい。各処理の処理主体は限定されるものではなく、矛盾の無い範囲で、各装置の処理を他の装置が実行してもよい。

[0107] 各実施形態に記載した事項は相互に組み合わせることが可能である。また、請求の範囲に記載した独立請求項及び従属請求項は、引用形式に関わらず全てのあらゆる組み合わせにおいて、相互に組み合わせることが可能である。さらに、請求の範囲には他の 2 以上のクレームを引用するクレームを記載

する形式（マルチクレーム形式）を用いているが、これに限るものではない。マルチクレームを少なくとも一つ引用するマルチクレーム（マルチマルチクレーム）を記載する形式を用いて記載してもよい。

[0108] 以上の実施形態で説明した事項は、以下の付記のようにも記載され得る。

（付記 1）

グループ内における複数のユーザにより作成された複数の文書を記憶しておき、

記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出力し、

出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取得し、

選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、

抽出した前記詳細情報を出力する

処理をコンピュータに実行させるためのコンピュータプログラム。

（付記 2）

前記概要情報は、文書の一部を表すサムネイル又は抜粋を含む

付記 1 に記載のコンピュータプログラム。

（付記 3）

検索キーワード又は絞り込み条件を含む検索条件を取得し、

取得した前記検索条件を満たす前記概要情報を特定し、

特定した前記概要情報を含む前記一覧情報を出力する

付記 1 又は付記 2 に記載のコンピュータプログラム。

（付記 4）

前記文書と、前記文書に関連する構造化データとを対応付けて取得し、

取得した前記文書及び前記構造化データを対応付けて記憶し、

前記構造化データに関する情報を前記詳細情報として出力する

付記 1 から付記 3 のいずれか 1 つに記載のコンピュータプログラム。

（付記 5）

前記詳細情報を含む訓練データを取得し、

取得した前記訓練データに基づいて所定の学習モデルを生成する

付記 1 から付記 4 のいずれか 1 つに記載のコンピュータプログラム。

(付記 6)

前記詳細情報は、物質に関する物質情報と、所定の物性を表す物性値とを含み、

前記物質情報と所定の物性を表す物性値とを含む前記訓練データに基づいて、物質に関する物質情報に応じた所定の物性を予測する前記学習モデルを生成する

付記 5 に記載のコンピュータプログラム。

(付記 7)

前記詳細情報を用いて学習モデルを生成するか否かの選択を取得し、

学習モデルを生成するとの選択を取得した場合、前記詳細情報を含む訓練データに基づいて所定の学習モデルを生成する

付記 1 から付記 6 のいずれか 1 つに記載のコンピュータプログラム。

(付記 8)

前記詳細情報は、物質に関する実験値、計算値及びプロセス情報の少なくとも 1 つを含む

付記 1 から付記 7 のいずれか 1 つに記載のコンピュータプログラム。

(付記 9)

前記詳細情報は、物質に関する実験値、計算値及びプロセス情報の少なくとも 1 つを含むファイルへのリンク情報を含む

付記 1 から付記 8 のいずれか 1 つに記載のコンピュータプログラム。

(付記 10)

前記文書は、所定形式に従い作成される表紙を備える

付記 1 から付記 9 のいずれか 1 つに記載のコンピュータプログラム。

(付記 11)

グループ内における複数のユーザにより作成された複数の文書を記憶して

おき、

記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出力し、

出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取得し、

選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、

抽出した前記詳細情報を出力する

情報処理方法。

(付記 1 2)

グループ内における複数のユーザにより作成された複数の文書を記憶装置に記憶し、

記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出力し、

出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取得し、

選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、

抽出した前記詳細情報を出力する

処理を実行する制御部を備える

情報処理装置。

(付記 1 3)

グループ内における複数のユーザにより作成された複数の文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を取得し、

取得した前記一覧情報を表示し、

表示した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を受け付け、

受け付けた前記概要情報の選択を出力し、

選択された概要情報に対応する文書の詳細を表す詳細情報を取得し、

取得した前記詳細情報を表示する

処理をコンピュータに実行させるためのコンピュータプログラム。

符号の説明

- [0109] 1 情報処理装置
- 1 1 制御部
- 1 2 記憶部
- 1 3 通信部
- 1 A 記録媒体
- 1 P プログラム
- 1 2 1 文書DB
- 1 2 2 学習モデル
- 2 端末装置
- 2 1 制御部
- 2 2 記憶部
- 2 3 通信部
- 2 4 表示部
- 2 5 操作部
- 2 A 記録媒体
- 2 P プログラム

請求の範囲

- [請求項1] グループ内における複数のユーザにより作成された複数の文書を記憶しておき、
- 記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出力し、
- 出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取得し、
- 選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、
- 抽出した前記詳細情報を出力する
- 処理をコンピュータに実行させるためのコンピュータプログラム。
- [請求項2] 前記概要情報は、文書の一部を表すサムネイル又は抜粋を含む
- 請求項1に記載のコンピュータプログラム。
- [請求項3] 検索キーワード又は絞り込み条件を含む検索条件を取得し、
- 取得した前記検索条件を満たす前記概要情報を特定し、
- 特定した前記概要情報を含む前記一覧情報を出力する
- 請求項1又は請求項2に記載のコンピュータプログラム。
- [請求項4] 前記文書と、前記文書に関連する構造化データとを対応付けて取得し、
- 取得した前記文書及び前記構造化データを対応付けて記憶し、
- 前記構造化データに関する情報を前記詳細情報として出力する
- 請求項1から請求項3のいずれか一項に記載のコンピュータプログラム。
- [請求項5] 前記詳細情報を含む訓練データを取得し、
- 取得した前記訓練データに基づいて所定の学習モデルを生成する
- 請求項1から請求項4のいずれか一項に記載のコンピュータプログラム。
- [請求項6] 前記詳細情報は、物質に関する物質情報と、所定の物性を表す物性

値とを含み、

前記物質情報と所定の物性を表す物性値とを含む前記訓練データに基づいて、物質に関する物質情報に応じた所定の物性を予測する前記学習モデルを生成する

請求項 5 に記載のコンピュータプログラム。

[請求項7] 前記詳細情報を用いて学習モデルを生成するか否かの選択を取得し、

学習モデルを生成するとの選択を取得した場合、前記詳細情報を含む訓練データに基づいて所定の学習モデルを生成する

請求項 1 から請求項 6 のいずれか一項に記載のコンピュータプログラム。

[請求項8] 前記詳細情報は、物質に関する実験値、計算値及びプロセス情報の少なくとも 1 つを含む

請求項 1 から請求項 7 のいずれか一項に記載のコンピュータプログラム。

[請求項9] 前記詳細情報は、物質に関する実験値、計算値及びプロセス情報の少なくとも 1 つを含むファイルへのリンク情報を含む

請求項 1 から請求項 8 のいずれか一項に記載のコンピュータプログラム。

[請求項10] 前記文書は、所定形式に従い作成される表紙を備える

請求項 1 から請求項 9 のいずれか一項に記載のコンピュータプログラム。

[請求項11] グループ内における複数のユーザにより作成された複数の文書を記憶しておき、

記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出し、

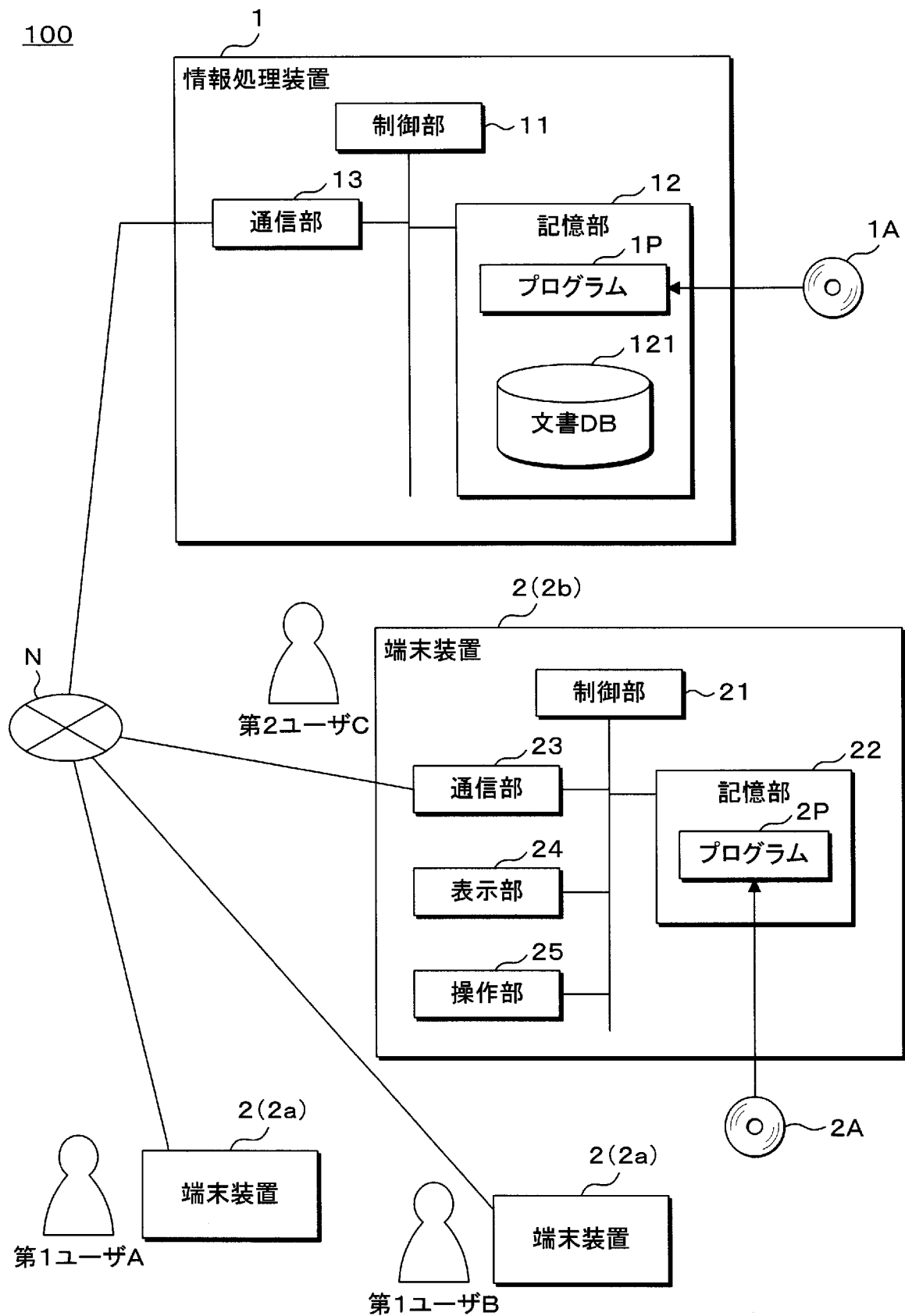
出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取得し、

選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、
抽出した前記詳細情報を出力する
情報処理方法。

[請求項12] グループ内における複数のユーザにより作成された複数の文書を記憶装置に記憶し、
記憶した各文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を出力し、
出力した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を取得し、
選択された概要情報に対応する文書の詳細を表す詳細情報を抽出し、
抽出した前記詳細情報を出力する
処理を実行する制御部を備える
情報処理装置。

[請求項13] グループ内における複数のユーザにより作成された複数の文書の概要を表す概要情報を選択可能に一覧で示す一覧情報を取得し、
取得した前記一覧情報を表示し、
表示した前記一覧情報における複数の概要情報のうちのいずれかの概要情報の選択を受け付け、
受け付けた前記概要情報の選択を出力し、
選択された概要情報に対応する文書の詳細を表す詳細情報を取得し、
取得した前記詳細情報を表示する
処理をコンピュータに実行させるためのコンピュータプログラム。

[図1]



[図2]

121

データID	文書情報			まとめファイル情報			実験ファイル情報		
	文書ID	文書ファイル	更新年月日	まとめファイルID	まとめファイル	リンク情報	実験ファイルID	実験ファイル	リンク情報
1	***	***	2023/1/1	x1	***	***	y1	***	***
2	***	***	***	x2	***	***	y2	***	***
...	...	...	...	...	...	***	***	***	***

〔図3〕

月報(2023年1月)

作成者	特許 太郎
研究所	第1研究所
研究グループ	第1グループ
チーム	〇〇チーム
作成者連絡先	<u>****@xxx.co.jp</u>
作成年月日	〇年〇月〇日

まとめファイル保存場所

実験1まとめ	/***/***/***/***/
実験2まとめ	/***/***/***/***/

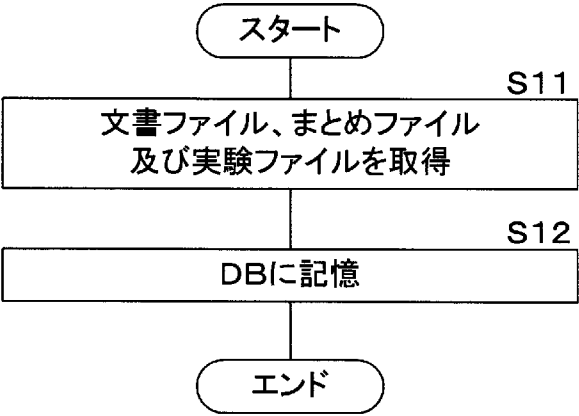
実験ファイル保存場所

プロセス条件	/***/***/***/***
分析データ	/***/***/***/***

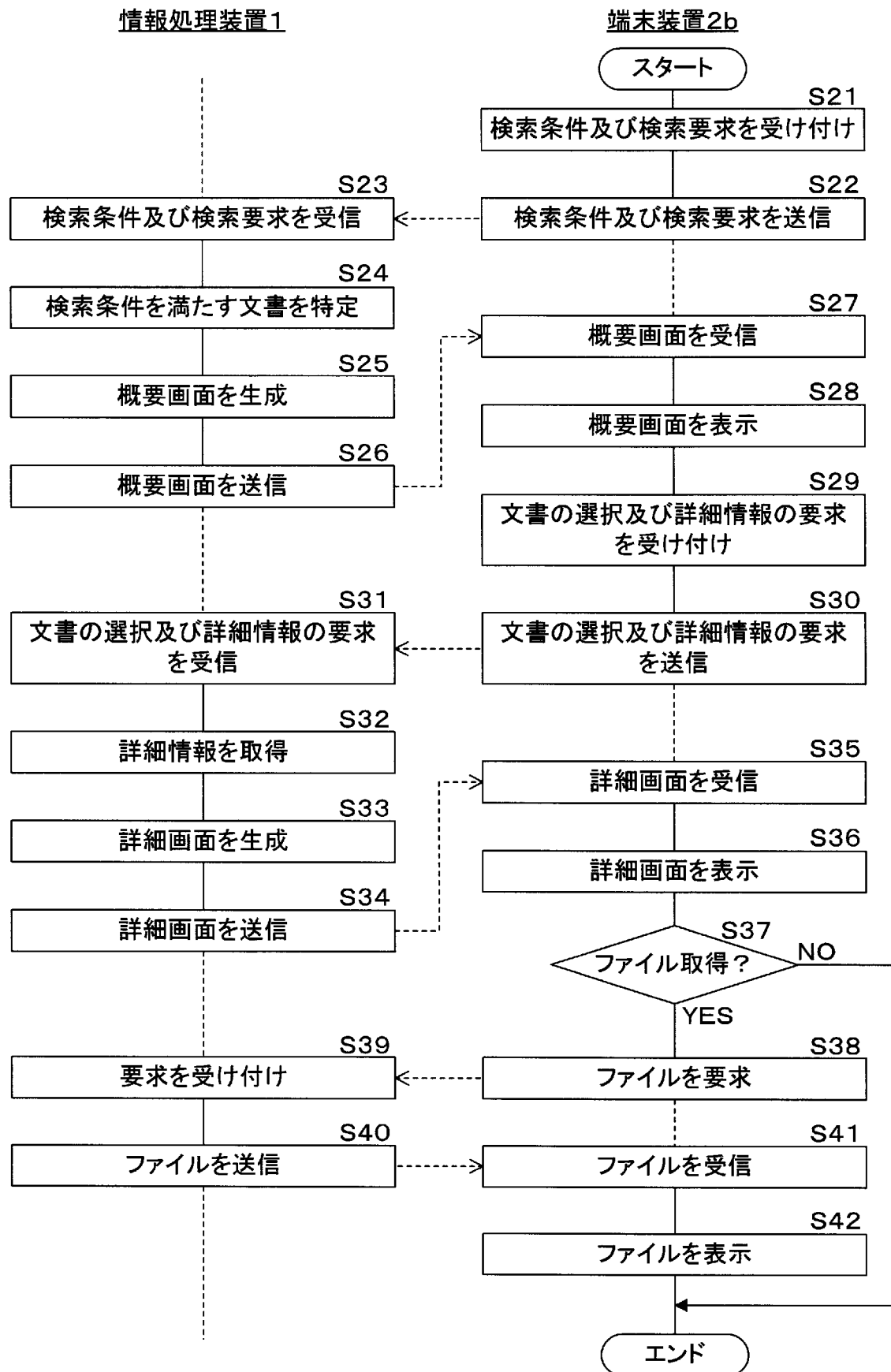
[図4]

サンプルID	注目物性 実験値1	物質情報				
		実験値2	計算値1	計算値2	プロセス値1	プロセス値2
a001	***	***	***	***	***	***
a002	***	***	***	***	***	***
...	...	...	...	...	...	...

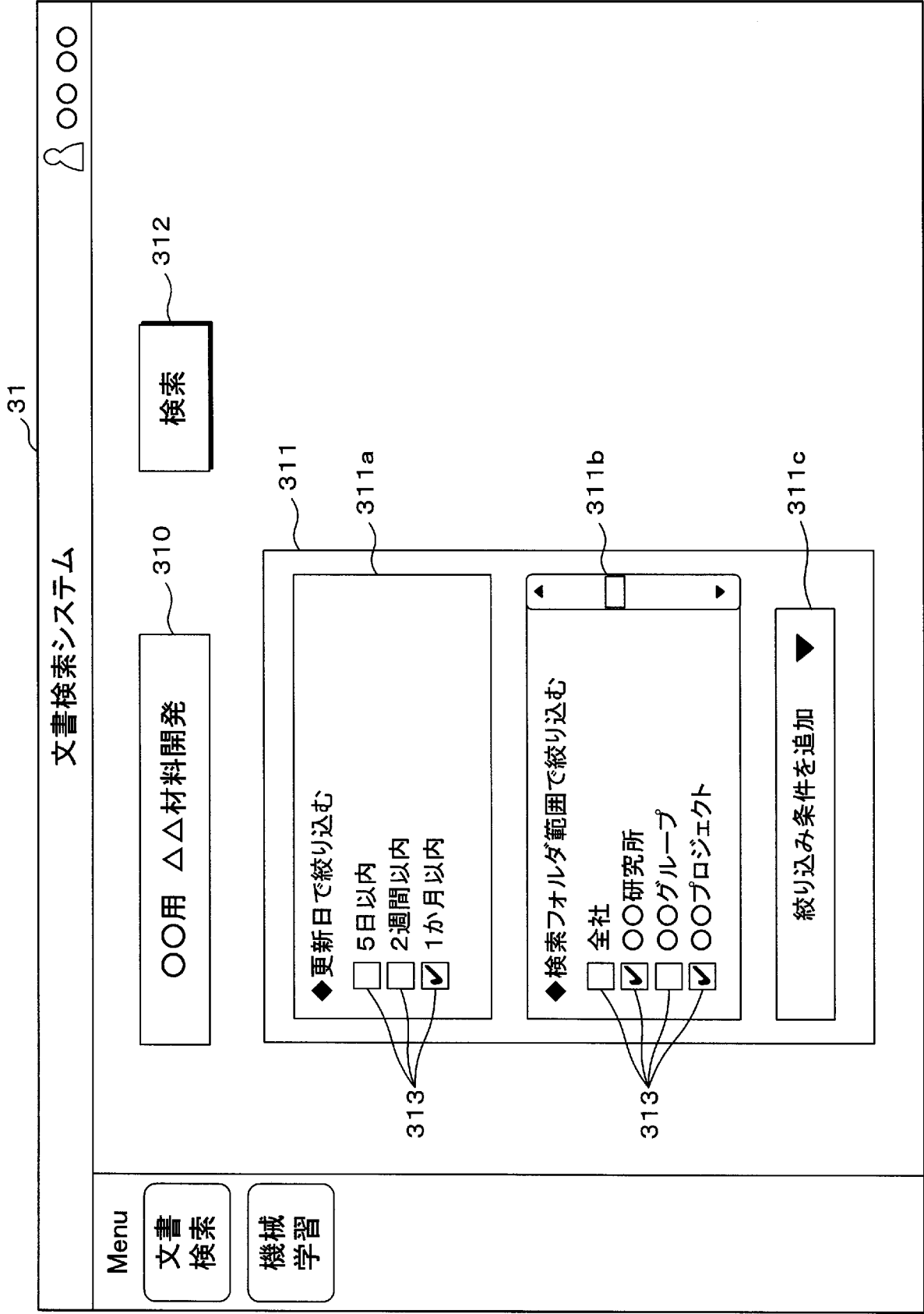
[図5]



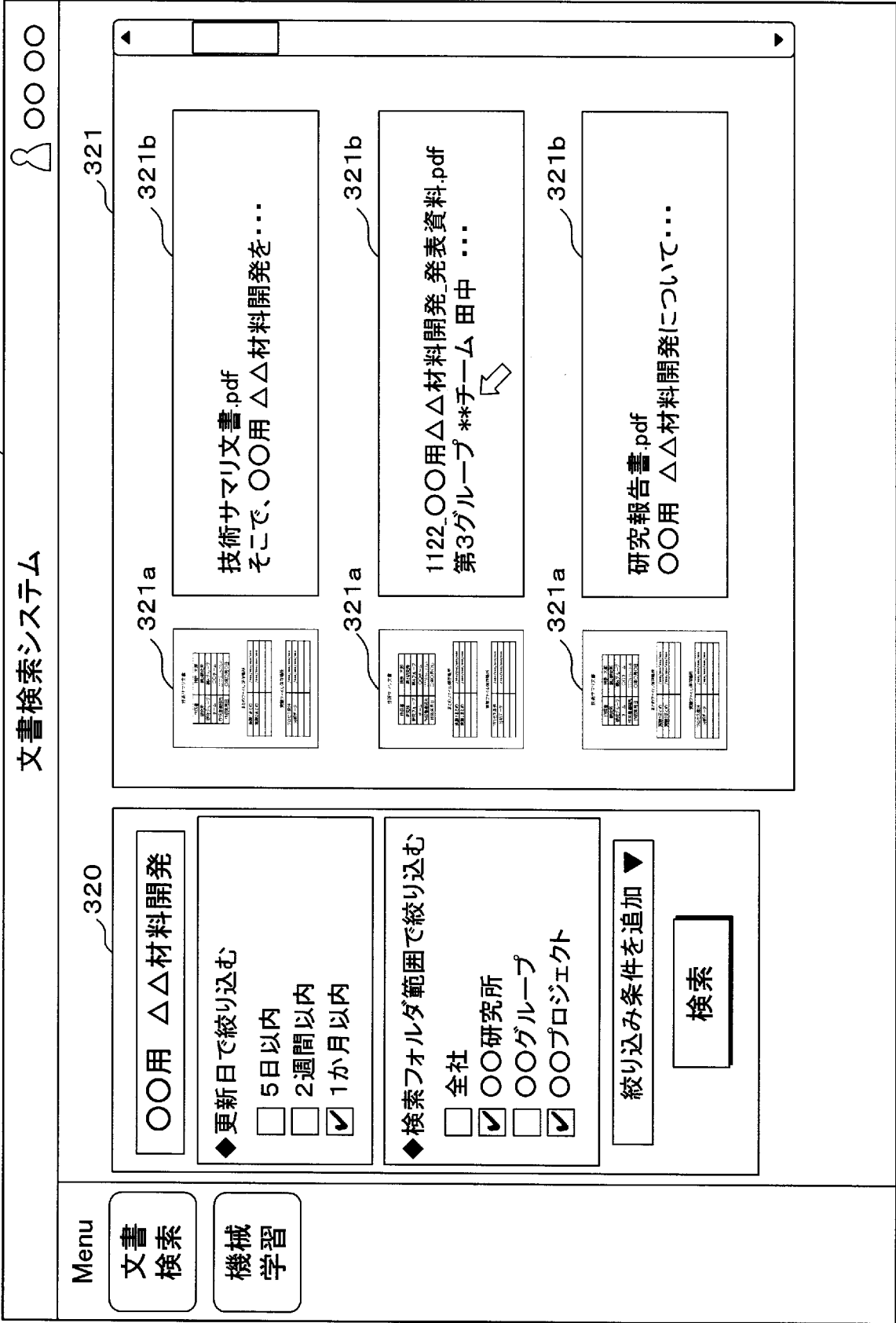
[図6]



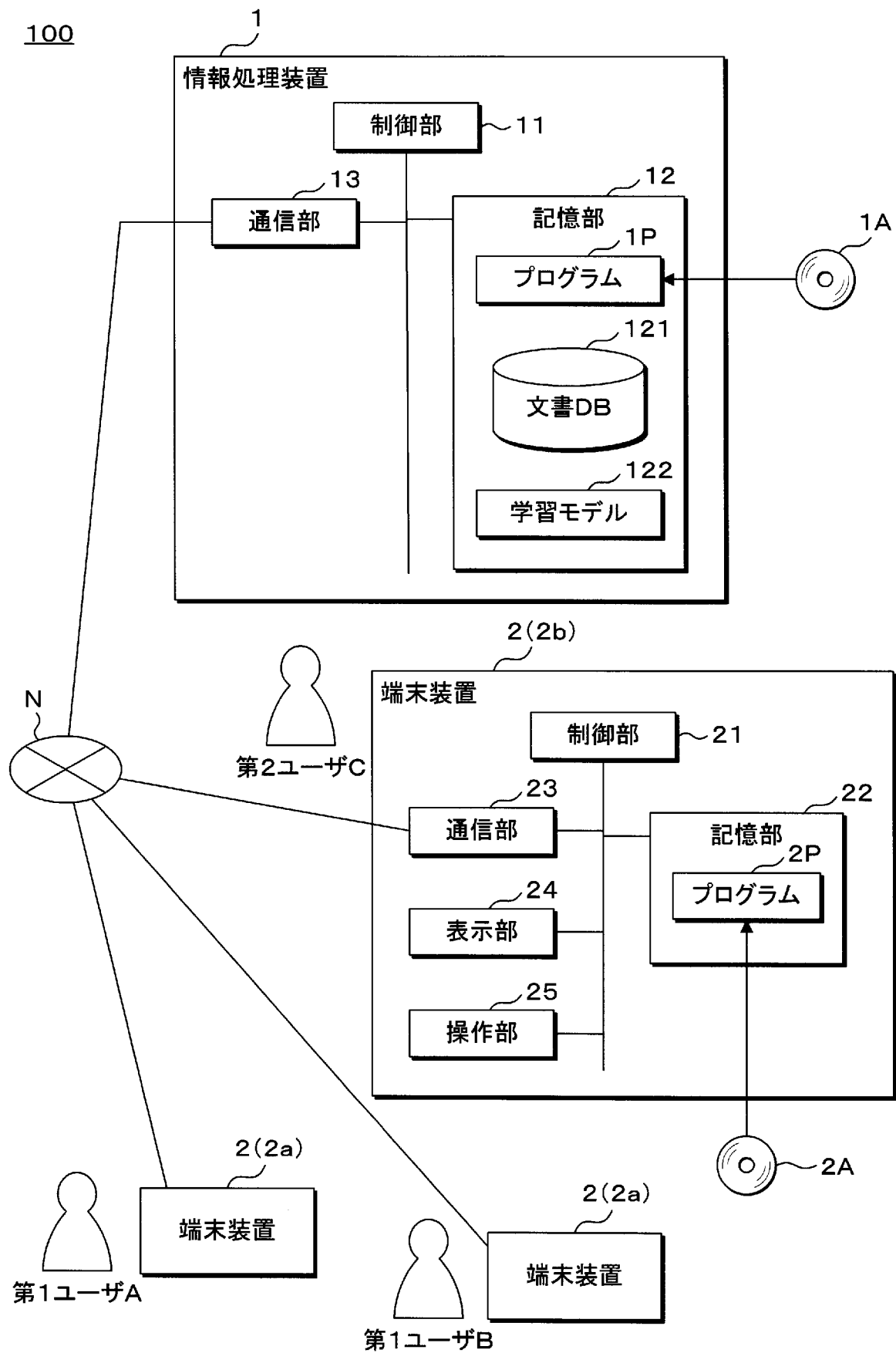
[図7]



[図8]



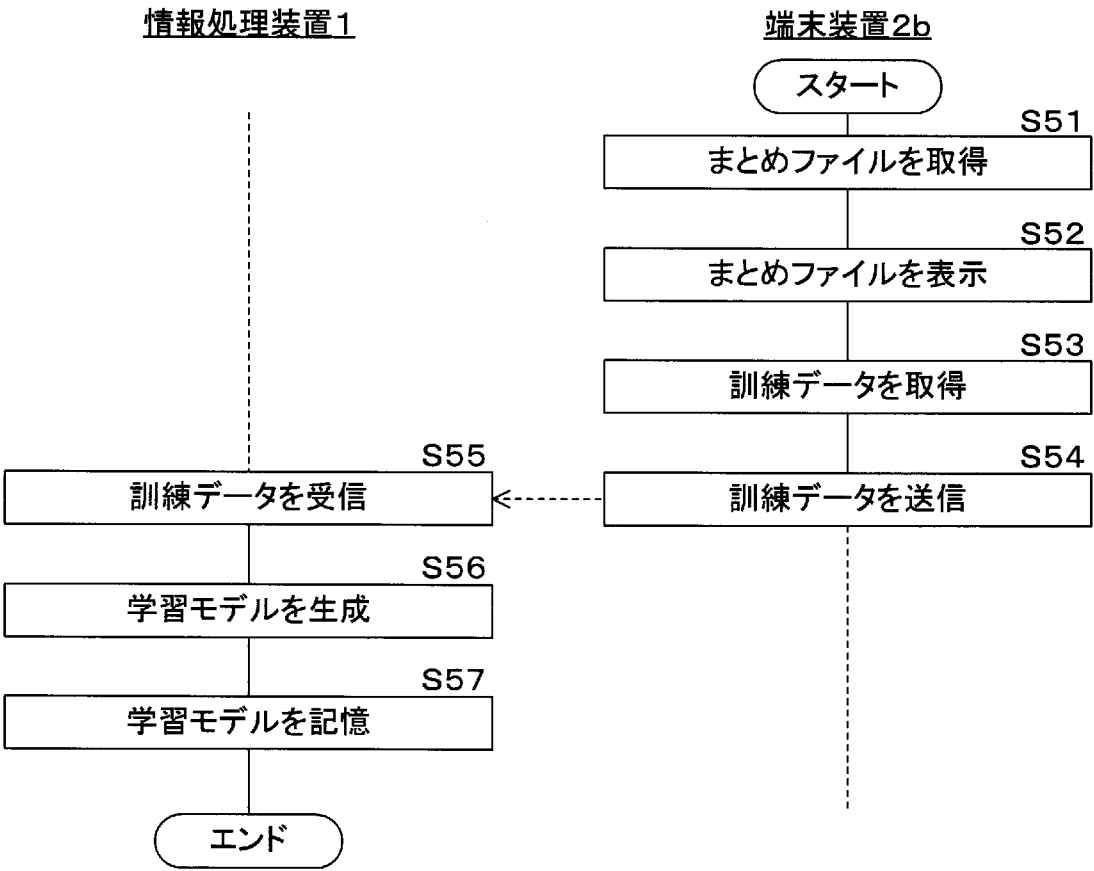
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2024/011851

A. CLASSIFICATION OF SUBJECT MATTER G06F 16/34 (2019.01)i; G06N 20/00 (2019.01)i FI: G06F16/34; G06N20/00 130 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F16/00-16/958; G06N20/00 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Published examined utility model applications of Japan 1922-1996 Published unexamined utility model applications of Japan 1971-2024 Registered utility model specifications of Japan 1996-2024 Published registered utility model applications of Japan 1994-2024 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2011-103075 A (THE KANSAI ELECTRIC POWER CO., INC.) 26 May 2011 (2011-05-26) paragraphs [0001], [0017]-[0021], [0065]	1-3, 8-13
Y		5-7
A		4
Y	森川幸治, 無機材料へのマテリアルズ インフォマティクスの取組みと課題, 人工知能, May 2019, vol. 34, no. 3, pp. 364-369, (MORIKAWA, Koji. Application of Materials Informatics to Inorganic Compounds. Journal of the Japanese Society for Artificial Intelligence.) In particular, "2.3 Construction of predictors", "2.4 Exploration of materials", and "3.3 Extraction of data from papers"	5-7
A		1-4, 8-13
A	US 2011/0313992 A1 (MICROSOFT CORPORATION) 22 December 2011 (2011-12-22) entire text, all drawings	1-13
☒ 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 “D” document cited by the applicant in the international application “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 29 May 2024		Date of mailing of the international search report 11 June 2024
Name and mailing address of the ISA/JP Japan Patent Office (ISA/JP) 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915 Japan		Authorized officer Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2018-173681 A (NOMURA RESEARCH INSTITUTE, LTD.) 08 November 2018 (2018-11-08) entire text, all drawings	1-13
A	JP 11-25092 A (JUST SYSTEMS CORP.) 29 January 1999 (1999-01-29) entire text, all drawings	1-13
A	JP 2007-26220 A (JAPAN WATER AGENCY, MIZUSHIGEN KYOKAI, TOTAL TECHNO JAPAN:KK) 01 February 2007 (2007-02-01) entire text, all drawings	1-13

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Claims are classified into the following two inventions.

(Invention 1) Invention in claims 1-4, 8-13, and part of claims 5-7 referring to claim 4

Claims 1-3 lack novelty in light of document 1 (see paragraphs [0001], [0017]-[0021], [0065]), and thus do not have a special technical feature. However, claim 4 dependent on claim 1 has a special technical feature and a part of claims 5-7 referring to claim 4 also has the same special technical feature as claim 4. Thus, claims 1-4 and the part of claims 5-7 referring to claim 4 are classified as invention 1.

(Invention 2) Part of claims 5-7 that is not classified as invention 1

It cannot be said that the part in claim 5 has a special technical feature identical or corresponding to that of claim 4 classified as invention 1.

Also, claims 5 are dependent on claim 1 classified as invention 1, but has little technical relevance to the technical feature of claim 5, added to claim 1. Accordingly, the part in claim 5 is not considered to be inventively related to claim 1.

This is also true with respect to the part in claims 6 and 7.

Furthermore, the part in claims 5-7 are not substantially identical to or similarly closely related to any of the claims classified as invention 1.

Accordingly, the above part of claims 5-7 cannot be classified as invention 1.

In addition, the above part of claims 5-7 is classified as invention 2.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2024/011851

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	2011-103075	A	26 May 2011	(Family: none)	
US	2011/0313992	A1	22 December 2011	(Family: none)	
JP	2018-173681	A	08 November 2018	(Family: none)	
JP	11-25092	A	29 January 1999	(Family: none)	
JP	2007-26220	A	01 February 2007	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC））

G06F 16/34(2019.01)i; G06N 20/00(2019.01)i

FI: G06F16/34; G06N20/00 130

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

G06F16/00-16/958; G06N20/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報

1922 - 1996年

日本国公開実用新案公報

1971 - 2024年

日本国実用新案登録公報

1996 - 2024年

日本国登録実用新案公報

1994 - 2024年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2011-103075 A（関西電力株式会社）26.05.2011（2011 - 05 - 26） 段落0001, 0017-0021, 0065	1-3, 8-13
Y		5-7
A		4
Y	森川幸治，無機材料へのマテリアルズインフォマティクスの取組みと課題，人工知能，2019.05，第34巻，第3号，pp. 364 - 369 特に、「2・3 予測器の構築」、「2・4材料の探索」、「3・3 論文からのデータ抽出」	5-7
A		1-4, 8-13
A	US 2011/0313992 A1（MICROSOFT CORPORATION）22.12.2011（2011 - 12 - 22） 全文全文図	1-13
A	JP 2018-173681 A（株式会社野村総合研究所）08.11.2018（2018 - 11 - 08） 全文全文図	1-13

☒ C欄の続きにも文献が列挙されている。

☒ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

“A” 特に関連のある文献ではなく、一般的技术水準を示すもの

“D” 国際出願で出願人が先行技術文献として記載した文献

“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

“O” 口頭による開示、使用、展示等に言及する文献

“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献

“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの

“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

“&” 同一パテントファミリー文献

国際調査を完了した日

29.05.2024

国際調査報告の発送日

11.06.2024

名称及びあて先

日本国特許庁(ISA/JP)

〒100-8915

日本国

東京都千代田区霞が関三丁目4番3号

権限のある職員（特許庁審査官）

早川 学 5N 9652

電話番号 03-3581-1101 内線 3519

C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 11-25092 A (株式会社ジャストシステム) 29.01.1999 (1999 - 01 - 29) 全文全図	1-13
A	JP 2007-26220 A (独立行政法人水資源機構、財団法人水資源協会、株式会社トータル・テクノ・ジャパン) 01.02.2007 (2007 - 02 - 01) 全文全図	1-13

第Ⅲ欄 発明の単一性が欠如しているときの意見（第1ページの3の続き）

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

請求の範囲は、以下の2つの発明に区分される。

（発明1）請求項1-4，8-13に係る発明、及び請求項5-7のうち請求項4を引用する部分

請求項1-3は、文献1（段落0001，0017-0021，0065参照。）により新規性が欠如しているため、特別な技術的特徴を有しない。しかしながら、請求項1の従属請求項である請求項4は、特別な技術的特徴を有しており、請求項5-7のうち、請求項4を引用している部分も、これと同一の特別な技術的特徴を有している。したがって、請求項1-4，及び請求項5-7のうち請求項4を引用する部分を発明1に区分する。

（発明2）請求項5-7のうち発明1に区分されない部分

請求項5の上記部分は、発明1に区分された請求項4と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項5は、発明1に区分された請求項1の従属請求項であるが、請求項5において請求項1に対して追加された技術的特徴と、技術的関連性が低い。このため、請求項5の上記部分は、請求項1に対して発明の連関を有しているとは認められない。

請求項6，7の上記部分についても同様である。

さらに、請求項5-7の上記部分は、発明1に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項5-7の上記部分は発明1に区分できない。

そして、請求項5-7の上記部分は、発明2に区分する。

1. ☒ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求項について作成した。
2. ☐ 追加調査手数料を要求するまでもなく、すべての調査可能な請求項について調査することができたので、追加調査手数料の納付を求めなかった。
3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったので、この国際調査報告は、手数料の納付のあった次の請求項のみについて作成した。
4. ☐ 出願人が必要な追加調査手数料を期間内に納付しなかったので、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求項について作成した。

追加調査手数料の異議の
申立てに関する注意

- ☐ 追加調査手数料及び、該当する場合には、異議申立手数料の納付と共に、出願人から異議申立てがあった。
- ☐ 追加調査手数料の納付と共に出願人から異議申立てがあったが、異議申立手数料が納付命令書に示した期間内に支払われなかった。
- ☒ 追加調査手数料の納付はあったが、異議申立てはなかった。

国際出願番号
PCT/JP2024/011851

様式 PCT/ISA/210 (パテントファミリー用別紙) (2022年7月)