

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5927809号  
(P5927809)

(45) 発行日 平成28年6月1日(2016.6.1)

(24) 登録日 平成28年5月13日(2016.5.13)

(51) Int.Cl. F I  
**G06Q 10/06 (2012.01)** G06Q 10/06  
**G06Q 50/10 (2012.01)** G06Q 50/10 100

請求項の数 24 (全 64 頁)

(21) 出願番号 特願2011-186447 (P2011-186447)  
 (22) 出願日 平成23年8月29日(2011.8.29)  
 (65) 公開番号 特開2012-48724 (P2012-48724A)  
 (43) 公開日 平成24年3月8日(2012.3.8)  
 審査請求日 平成26年7月11日(2014.7.11)  
 (31) 優先権主張番号 12/871,720  
 (32) 優先日 平成22年8月30日(2010.8.30)  
 (33) 優先権主張国 米国 (US)

(73) 特許権者 000006747  
 株式会社リコー  
 東京都大田区中馬込1丁目3番6号  
 (74) 代理人 100070150  
 弁理士 伊東 忠彦  
 (72) 発明者 ベルナ エロール  
 アメリカ合衆国, カリフォルニア 940  
 25-7054, メンロ・パーク, サンド  
 ・ヒル・ロード 2882番, スイート  
 115 リコー イノベーション インク  
 内

最終頁に続く

(54) 【発明の名称】 タスク値付け技術

(57) 【特許請求の範囲】

【請求項1】

処理システムによって、実行されるべきタスクに関する入力情報を受領する段階と、  
 前記処理システムによって、前記入力情報を解析して前記タスクに関する前記入力情報の一つまたは複数の属性を決定する段階であって、前記入力情報の前記一つまたは複数の属性は、前記入力情報の型、前記入力情報の内容、前記入力情報の複雑さまたは前記入力情報のコンテキストを含む、段階と、

前記処理システムによって、前記一つまたは複数の属性および前記タスクの実行の値付けに影響する情報に基づいて、前記タスクの価格を下げるよう前記入力情報を修正して修正された入力情報を生成する段階であって、タスクの価格を前記タスクの前記一つまたは複数の属性および前記タスクの実行の値付けに影響する情報に基づいて決定する一つまたは複数の規則の集合が与えられており、前記タスクの実行の値付けに影響する情報は、前記タスクの完了の所望される時間、前記タスクについて所望される精度のレベル、前記タスクを実行する位置、前記タスクを実行するために必要とされるスキル、前記タスクを実行するために提供されるリソースまたは前記タスクの難しさのレベルのうちの一つまたは複数を含む、段階と、

前記処理システムによって、修正された入力情報についての前記タスクの第一の価格を決定する段階と、

前記処理システムによって、前記タスクが前記第一の価格での実行のために分配されるべき作業システムを決定する段階とを含む、

10

20

方法。

【請求項 2】

請求項 1 記載の方法であって、さらに：

前記タスクを実行する複数のサブタスクを決定する段階を含む、

方法。

【請求項 3】

請求項 1 記載の方法であって、前記タスクが全体タスクについて決定される複数のタスクを含む、方法。

【請求項 4】

請求項 3 記載の方法であって、前記タスクの前記第一の価格が、前記全体タスクの目標価格に基づいて決定される、方法。

10

【請求項 5】

請求項 1 記載の方法であって、さらに：

前記処理システムによって、前記タスクの完了から生じる結果を受領する段階と、

前記処理システムによって、前記結果の精度のレベルが前記タスクについての期待される精度値より低いことを判別する段階と、

前記処理システムによって、前記結果の精度のレベルについての情報を、前記修正を実行するセグメント分割器サブシステムに通信する段階と、

前記処理システムによって、前記セグメント分割器サブシステムから前記修正された入力情報を受領する段階とを含む、

20

方法。

【請求項 6】

請求項 5 記載の方法であって、前記修正された入力情報が、前記入力情報からの、コンピュータによって処理可能な情報を含む第一の部分と、前記入力情報からの、人間によって処理可能な情報を含む第二の部分とを含む、方法。

【請求項 7】

メモリおよび該メモリに結合されたプロセッサを有するシステムであって、前記プロセッサは：

実行されるべきタスクに関する入力情報を受領する段階と、

前記入力情報を解析して前記入力情報の一つまたは複数の属性を決定する段階であって、前記入力情報の前記一つまたは複数の属性は、前記入力情報の型、前記入力情報の内容、前記入力情報の複雑さまたは前記入力情報のコンテキストを含む、段階と、

30

前記一つまたは複数の属性および前記タスクの実行の値付けに影響する情報に基づいて、前記タスクの価格を下げるよう前記入力情報を修正する段階であって、タスクの価格を前記タスクの前記一つまたは複数の属性および前記タスクの実行の値付けに影響する情報に基づいて決定する一つまたは複数の規則の集合が与えられており、前記タスクの実行の値付けに影響する情報は、前記タスクの完了の所望される時間、前記タスクについて所望される精度のレベル、前記タスクを実行する位置、前記タスクを実行するために必要とされるスキル、前記タスクを実行するために提供されるリソースまたは前記タスクの難しさのレベルのうちの一つまたは複数を含む、段階と、

40

修正された入力情報について前記タスクの第一の価格を決定する段階と、

前記第一の価格での実行のために前記タスクが分配されるべき作業システムを決定する段階とを実行するよう構成されている、

システム。

【請求項 8】

請求項 7 記載のシステムであって、前記プロセッサはさらに：

少なくとも部分的には前記入力情報の解析に基づいて、前記タスクを実行するための複数のサブタスクを決定する段階と、

前記複数のサブタスクからの第一のサブタスクについて：

前記第一のサブタスクに関連付けられた前記入力情報の部分を解析する段階と、

50

前記第一のサブタスクに関連付けられた前記入力情報の前記部分および前記タスクの前記第一の価格に部分的に基づいて、前記第一のサブタスクの価格を決定する段階とを  
 実行するよう構成されている、  
 システム。

【請求項 9】

請求項 7 記載のシステムであって、前記タスクが全体タスクについて決定される複数のタスクの一部である、システム。

【請求項 10】

請求項 9 記載のシステムであって、前記プロセッサがさらに、前記タスクの前記第一の価格を、前記全体タスクの目標価格に基づいて決定するよう構成されている、システム。

10

【請求項 11】

請求項 7 記載のシステムであって、前記プロセッサがさらに：  
 前記タスクの完了から生じる結果を受領する段階と、  
 前記結果の精度のレベルが前記タスクについての期待される精度値より低いことを判別する段階と、  
 前記結果の精度のレベルについての情報を、前記修正を実行するセグメント分割器サブシステムに通信する段階と、  
 前記セグメント分割器サブシステムから前記修正された入力情報を受領する段階と  
 を実行するよう構成されている、  
 システム。

20

【請求項 12】

請求項 11 記載のシステムであって、前記修正された入力情報が、前記入力情報からの、コンピュータによって処理可能な情報を含む第一の部分と、前記入力情報からの、人間によって処理可能な情報を含む第二の部分とを含む、システム。

【請求項 13】

請求項 7 記載のシステムであって、前記プロセッサがさらに：  
 第一のセグメントおよび第二のセグメントを受領する段階であって、前記第一および第二のセグメントのそれぞれは前記入力情報の一部を表し、前記第一のセグメントは前記第二のセグメントとは異なる、段階と、  
 一つまたは複数の規則を使って、前記第一のセグメントの価格および前記第二のセグメントの価格を決定する段階とを実行するよう構成されている、  
 システム。

30

【請求項 14】

請求項 7 記載のシステムであって、前記プロセッサがさらに：  
 前記タスクの前記第一の価格を前記タスクの目標価格と比較する段階と、  
 前記タスクの前記第一の価格が前記目標価格より高い場合、前記入力情報の前記修正を実行するようセグメント分割器サブシステムに命令する段階と、  
 前記命令に応答して前記修正された入力情報を受領する段階とを実行するよう構成されている、  
 システム。

40

【請求項 15】

請求項 7 記載のシステムであって、前記第一の価格が目標価格の所定の許容範囲内である、システム。

【請求項 16】

タスクの価格を決定するようプロセッサを制御する複数の命令を記憶しているコンピュータ可読記憶媒体であって、前記複数の命令は：  
 前記プロセッサに、実行されるべきタスクに関する入力情報を受領させる命令と、  
 前記プロセッサに、前記入力情報を解析して前記入力情報の一つまたは複数の属性を決定させる命令であって、前記入力情報の前記一つまたは複数の属性は、前記入力情報の型、前記入力情報の内容、前記入力情報の複雑さまたは前記入力情報のコンテキストを含む

50

、命令と、

前記プロセッサに、前記一つまたは複数の属性および前記タスクの実行の値付けに影響する情報に基づいて、前記タスクの価格を下げるよう前記入力情報を修正して修正された入力情報を生成させる命令であって、タスクの価格を前記タスクの前記一つまたは複数の属性および前記タスクの実行の値付けに影響する情報に基づいて決定する一つまたは複数の規則の集合が与えられており、前記タスクの実行の値付けに影響する情報は、前記タスクの完了の所望される時間、前記タスクについて所望される精度のレベル、前記タスクを実行する位置、前記タスクを実行するために必要とされるスキル、前記タスクを実行するために提供されるリソースまたは前記タスクの難しさのレベルのうちの一つまたは複数を含む、命令と、

10

前記プロセッサに、前記タスクの第一の価格を決定させる命令と、

前記プロセッサに、前記第一の価格での実行のために前記タスクが分配されるべき作業システムを決定させる命令とを含む、  
コンピュータ可読記憶媒体。

【請求項 17】

請求項 16 記載のコンピュータ可読記憶媒体であって、前記複数の命令はさらに：

前記プロセッサに、前記タスクを実行する複数のサブタスクを決定させる命令を含む、  
コンピュータ可読記憶媒体。

【請求項 18】

請求項 16 記載のコンピュータ可読記憶媒体であって、前記タスクが全体タスクについて決定される複数のタスクの一部である、コンピュータ可読記憶媒体。

20

【請求項 19】

請求項 18 記載のコンピュータ可読記憶媒体であって、前記複数の命令がさらに、前記プロセッサに、前記タスクの前記第一の価格を、前記全体タスクの目標価格に基づいて決定させる命令を含む、コンピュータ可読記憶媒体。

【請求項 20】

請求項 16 記載のコンピュータ可読記憶媒体であって、前記複数の命令がさらに：

前記プロセッサに、前記タスクの完了から生じる結果を受領させる命令と、  
前記プロセッサに、前記結果の精度のレベルが前記タスクについての期待される精度値より低いことを判別させる命令と、

30

前記プロセッサに、前記結果の精度のレベルについての情報を前記セグメント分割器サブシステムに通信させる命令と、

前記プロセッサに、前記セグメント分割器サブシステムから前記修正された入力情報を受領させる命令とを含む、  
コンピュータ可読記憶媒体。

【請求項 21】

請求項 20 記載のコンピュータ可読記憶媒体であって、前記修正された入力情報が、前記入力情報からの、コンピュータによって処理可能な情報を含む第一の部分と、前記入力情報からの、人間によって処理可能な情報を含む第二の部分とを含む、コンピュータ可読記憶媒体。

40

【請求項 22】

請求項 16 記載のコンピュータ可読記憶媒体であって、前記複数の命令がさらに：

前記プロセッサに第一のセグメントおよび第二のセグメントを受領させる命令であって、前記第一および第二のセグメントのそれぞれは前記入力情報の一部を表し、前記第一のセグメントは前記第二のセグメントとは異なる、命令と、

前記プロセッサに、一つまたは複数の規則を使って、前記第一のセグメントの第一の価格および前記第二のセグメントの第二の価格を決定させる命令とを含む、  
コンピュータ可読記憶媒体。

【請求項 23】

請求項 16 記載のコンピュータ可読記憶媒体であって、前記第一の価格が目標価格の所

50

定の許容範囲内である、コンピュータ可読記憶媒体。

【請求項 2 4】

請求項 1 6 記載のコンピュータ可読記憶媒体であって、前記入力情報がテキスト、画像、グラフィック、ビデオ情報またはオーディオ情報のうちの少なくとも一つを含む、コンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願への相互参照 本願は、2010年8月30日に本願と同時に出願された「コンテンツ・プライバシー保護のためのマイクロタスクを生成する技術」という名称の米国特許出願第12/871,665号（代理人整理番号015358-012900US）に関係し、その全内容をあらゆる目的において参照により援用する。

10

【0002】

本発明の諸実施形態は、データ処理に関し、より詳細には、入力情報およびタスクについて決定される一つまたは複数の規則に基づいてタスクに値付けするための技法に関する。

【背景技術】

【0003】

コンピュータおよび人工知能（AI: artificial intelligence）技術の発達にもかかわらず、実行されるには、あるいは効率的もしくは正確に実行されるには人間の知性によるしかないタスクがある。そのようなタスクの例は、写真やビデオにおけるオブジェクト認識、手書き文字認識、手書き文字からテキストへの変換、翻訳、転写などが含まれる。これらのタスクについては、自動化技法を使ってタスクを実行することから得られる正確さは、人間の知性を使って人間によってタスクが実行されるときに得られる正確さには遠く及ばない。

20

【0004】

その本来的な性質上、そのようなタスクは典型的には、完成させるために人間に与えられる。タスクを実行する人間は通例、タスクの実行について何らかの対価を与えられる。タスクがその意図される人間の作業者に分配される種々の方法がある。インターネットのような通信ネットワークの到来により、タスクが電子的に人間の作業者に分配されることを可能にするいくつかのオンライン・コミュニティが出てきている。インターネットの広範な到達範囲のため、そのようなタスクは今では、米国内の種々の地理的位置を含む多様な位置にいる人間の作業者に、またさらにはタスクを実行するための必要なスキルをもち、著しく低価格でタスクを実行しうる作業者がいる外国の作業者にさえも、電子的にアウトソーシングされうる。

30

【0005】

人間の知性を必要とするタスクを人間の作業者に送達し、タスクの実行の結果を回収するプロセスを指すのに、「マイクロ・アウトソーシング」の用語が使われることがある。たとえば、アマゾン（Amazon）は、人間の知性を必要とする作業（そのようなタスクは一般に人間知性タスク（human intelligence task）またはHITと称される）のためのオンライン市場を提供するアマゾン・メカニカル・ターク [機械仕掛けのトルコ人]（AMT: Amazon Mechanical Turk）と呼ばれるオンライン・コミュニティを提供している。このメカニカル・タークは、人間の作業員 [ワーカー] および作業要求者 [リクエスト] の市場を提供する、ウェブ・ベースのマイクロ・アウトソーシング・サービスを提供する。AMTによって提供されるアプリケーション・プログラミング・インターフェース（API）を使って、作業要求者は、タスク明細、タスク実行についての価格、タスク完了のための時間枠、タスクについての所望される品質、タスクを実行する人間の作業者の所望される位置などのようなHITのためのパラメータを指定してもよい。AMTの最も一般的な使用は、音声文字起こし、ブログ項目執筆および画像のタグ付けを含む。AMTのようなサービスは、作業要求者である企業に、HITを実行するための多様な、オンデマンドの労働人口をも

40

50

つ市場にプログラマ的にアクセスすることを可能にする。

【先行技術文献】

【非特許文献】

【0006】

【非特許文献1】Berna Erol et al., "HOTPAPER: Multimedia Interaction with Paper using Mobile Phones," ACM Multimedia Conference, 2008, Vancouver, British Columbia, Canada, pp.399-408

【発明の概要】

【発明が解決しようとする課題】

【0007】

マイクロ・アウトソーシングのもう一つの側面は、タスクの値付けである。現在のところ、AMTなどでのタスクはタスクを発生させる人/エンティティ(タスク発生者)によって値付けされる。しばしば、タスクについての価格はタスク発生者によって、タスクについての自分の知識および/またはタスクがどのくらいのコストがかかるべきかについての期待に基づいて、恣意的に決定される。タスク値付けの恣意的な性質のため、タスクについて提供される価格は十分な競争性がないことがある。しばしば、価格は人々が受け容れるには十分うまみがないことがある。その結果、タスクが応札や実行されないことになり、複数ラウンドの値付けし直しにつながり、貴重な時間および資源が浪費されることになる。タスクのさまざまな側面を考慮に入れ、タスクの性質に見合っている可能性が高い価格を提供する包括的な値付けシステムが必要とされている。

【課題を解決するための手段】

【0008】

本発明の諸実施形態は、顧客によって与えられる情報を使って実行されるべきタスクに値付けするための技法を提供する。タスクについての価格は、顧客によって与えられる入力情報のさまざまな属性ならびにそのタスクについて適用可能な一つまたは複数の規則に基づいて計算されてもよい。

【0009】

本発明の諸実施形態は、タスクに値付けする技法を提供する。本技法は、実行されるべきタスクのための入力情報を受領し、入力情報を解析して入力情報の一つまたは複数の属性を決定することを含む。いくつかの実施形態では、前記一つまたは複数の属性は、テキスト文書中の語数、オーディオ/ビデオ・コンテンツの長さ、入力情報の複雑さを含んでもよい。本技法はさらに、タスクについての値付けを決定するための一つまたは複数の規則の集合を決定し、入力情報の属性および規則の集合に基づいてタスクについての価格を決定することを含む。

【0010】

本発明のある種の実施形態は、タスクを値付けするためのシステムを提供する。システムはメモリおよび該メモリに結合されたプロセッサを有する。プロセッサは、実行されるべきタスクのための入力情報を受領し、入力情報を解析して入力情報の一つまたは複数の属性を決定するよう構成される。プロセッサは、タスクについての値付けを決定するための一つまたは複数の規則の集合を決定し、入力情報の属性および規則の集合に基づいてタスクについての価格を決定してもよい。いくつかの実施形態では、プロセッサは、タスクの完了から生じる結果を受領し、該結果の精度のレベルがそのタスクについての期待される精度値より下であることを判別し、該結果の精度のレベルについての情報をセグメント分割器サブシステムに通信するようさらに構成される。その後、プロセッサは、セグメント分割器サブシステムから修正された入力情報を受領し、修正された入力情報に基づいてタスクについての第二の価格を決定してもよい。

【0011】

本発明の他の諸実施形態は、タスクについての価格を決定するようプロセッサを制御するための複数の命令を記憶するコンピュータ可読記憶媒体を提供する。前記複数の命令は、前記プロセッサに、実行されるべきタスクのための入力情報を受領させる命令と、前記

10

20

30

40

50

プロセッサに、入力情報を解析して入力情報の一つまたは複数の属性を決定させる命令と、前記プロセッサに、タスクについての値付けを決定するための規則の集合を決定させる命令と、前記プロセッサに、入力情報の属性および規則の集合に基づいてタスクについての価格を決定させる命令とを含んでもよい。コンピュータ可読記憶媒体はまた、前記プロセッサに、タスクの完了から生じる結果を受領させる命令と、前記プロセッサに、前記結果の精度のレベルがそのタスクについての期待される精度値より下であることを判別させる命令とを記憶していてもよい。前記プロセッサに、前記結果の精度のレベルについての情報をセグメント分割器サブシステムに通信させる他の命令、前記プロセッサに、セグメント分割器サブシステムから修正された入力情報を受領させる命令、前記プロセッサに、修正された入力情報に基づいてタスクについての第二の価格を決定させる命令があつてもよい。

10

**【 0 0 1 2 】**

以下の詳細な説明は、付属の図面とともに、本発明の実施形態の性質および利点のよりよい理解を与えるであろう。

**【図面の簡単な説明】****【 0 0 1 3 】**

【図 1】本発明のある実施形態を組み込みうるシステムの概略図である。

【図 2】本発明のある実施形態に基づく、タスクのために受領された入力情報の内容のプライベートを保護しつつタスクを実行する高レベルの方法を記述する簡略化されたフローチャートである。

20

【図 3】本発明のある実施形態に基づく、入力情報をセグメント分割するために実行される処理を記述する簡略化されたフローチャートである。

【図 4】本発明のある実施形態に基づく、組み合わせられたセグメントを生成するために実行される処理を記述する簡略化されたフローチャートである。

【図 5】本発明のある実施形態に基づく、マイクロタスク生成器サブシステムによって実行される処理を示す簡略化されたフローチャートである。

【図 6】本発明のある実施形態に基づく、分配システムによって処理を示す簡略化されたフローチャートである。

【図 7】本発明のある実施形態に基づく、タスクに対応するマイクロタスクについて受領されたマイクロタスク生成物に基づいてタスクについての最終作業生成物を生成するために実行される処理を記述する簡略化されたフローチャートである。

30

【図 8 A】本発明のある実施形態に基づくさまざまな側面を記述する例である。

【図 8 B】本発明のある実施形態に基づくさまざまな側面を記述する例である。

【図 9】本発明のある実施形態に基づく、マイクロタスクのための価格を決定するためのシステムのブロック図である。

【図 1 0】本発明のある実施形態に基づく、タスクのための価格付けを決定するためのシステムのブロック図である。

【図 1 1】本発明のある実施形態に基づく手書き文字認識アプリケーションを示す図である。

【図 1 2】本発明のある実施形態に基づくレシート認識アプリケーションを示す図である。

40

【図 1 3】本発明のある実施形態に基づく名刺認識アプリケーションを示す図である。

【図 1 4 A】本発明のある実施形態に基づく描画変換アプリケーションを示す図である。

【図 1 4 B】本発明のある実施形態に基づく描画変換アプリケーションを示す図である。

【図 1 5】本発明のある実施形態を実施するために使用されうるコンピュータ・システムのブロック概略図である。

**【発明を実施するための形態】****【 0 0 1 4 】**

以下の記述では、説明の目的のために、本発明の実施形態の十全な理解を与えるために、個別的な詳細が述べられるが、本発明がこれらの個別的な詳細なしでも実施できること

50

は明白であろう。

【 0 0 1 5 】

本発明の諸実施形態は、タスクのための入力として使われる情報のプライバシーまたは秘匿性を保持しつつタスクを実行するための技法を提供する。ある実施形態では、タスクはより小さなタスク（サブタスクまたはマイクロタスクと呼ばれる）に分解される。各マイクロタスクについての入力情報は、そのタスクについて受領された入力情報に基づき、一般にその部分集合である。タスクについてのマイクロタスクの決定は、タスクに関連する制約条件が満足されるような仕方で行われる。たとえば、タスクについて、マイクロタスクは、リスク（たとえば、アウトソーシングの結果として入力情報のプライバシーまたは秘匿性が危殆化されることに関わるリスク）、品質制約条件（たとえば、タスクの実行から帰結する作業生成物の所望される品質）、コスト制約条件およびジョブに関連する他の制約条件に基づいて決定されてもよい。

10

【 0 0 1 6 】

本発明のある種の実施形態は、タスクに値付けする技法を提供する。ある実施形態では、本技法は、実行されるべきタスクのための入力情報を受領し、入力情報を解析して入力情報の一つまたは複数の属性を決定することを含む。いくつかの実施形態では、前記一つまたは複数の属性は、テキスト文書中の語数、オーディオ/ビデオ・コンテンツの長さ、入力情報の複雑さを含んでいてもよい。本技法はさらに、タスクについての値付けを決定するための一つまたは複数の規則の集合を決定し、入力情報の属性および規則の集合に基づいてタスクについての価格を決定することを含む。

20

【 0 0 1 7 】

図 1 は、本発明のある実施形態を組み込んでよいシステム 1 0 0 の概略図である。システム 1 0 0 は、一つまたは複数のタスク要求者システム 1 0 2 と、マイクロタスク管理システム（MMS: microtask management system）1 0 4 と、分配システム 1 0 6 と、アウトソーシングされたマイクロタスクを実行するために使用されうる一つまたは複数のシステム 1 0 8 および 1 1 0 とを含む複数のシステムを有する。図 1 に描かれるさまざまなシステムは、一つまたは複数の通信ネットワーク（図 1 には示さず）を介して互いに結合されていてもよい。これらの通信ネットワークは、同じ型であっても異なる型であってもよく、インターネット、イントラネット、ローカル・エリア・ネットワーク（LAN）、広域ネットワーク（WAN）、イーサネット（登録商標）・ネットワークまたはシステム間通信を可能にする任意の通信ネットワーク/インフラストラクチャーを含んでいてもよい。これらの通信ネットワークは、通信のための有線または無線のプロトコルを含む異なる通信プロトコルを使ってもよい。図 1 に描かれるシステム 1 0 0 は単に、本発明の教示を組み込むある実施形態の一例であり、請求項に記載される本発明の範囲を限定することを意図したものではない。

30

【 0 0 1 8 】

高レベルでは、MMS 1 0 4 はタスク要求者からタスク要求を受領する。タスク要求は、実行されるべきタスクと、そのタスクを実行するための入力として使われる情報とを同定してもよい。MMS 1 0 4 は、タスクについての入力情報のプライバシーを保持するためのステップを講じつつ、タスクの実行を容易にするよう構成される。ある実施形態では、MMS 1 0 4 は、実行されるべきタスクに対応する、アウトソーシングされるべき一組のタスク（サブタスクまたはマイクロタスクと称される）を決定し、マイクロタスクについて入力として使われるべき前記入力情報の部分を決定するよう構成される。タスクのマイクロタスクへの分割は、そのタスクについて受領される入力情報のプライバシーを保護するために（換言すれば、マイクロタスクをアウトソーシングする結果として入力情報のプライバシーまたは秘匿性が破られるリスクを下げるために）実行される。ある実施形態では、タスク要求は、そのタスクについての受け入れ可能なリスク閾値を指定してもよい。このリスク閾値は、タスクをマイクロタスクに細分するために考慮に入れられる。タスク要求はまた、タスクを実行する受け入れ可能なコスト、タスクを実行する結果として生成される出力についての所望される品質などといったタスクに関係する他の因子を指定し

40

50



てもよい。これらのさまざまな因子も、どのようにタスクをマイクロタスクに細分するかを決定するときに考慮されてもよい。次いで、MMS 104は、マイクロタスクを実行するプロバイダー114への分配またはアウトソーシングのために、マイクロタスクおよびそれぞれの入力情報を分配システムに転送する。

#### 【0019】

プロバイダー114は人間の作業員および/または自動化されたシステム(機械)110を含んでいてもよい。本願で使われるところの用語「作業員」は、タスクまたはマイクロタスクを実行する、人間の作業員または機械を指しうる。人間の作業員は、その作業員に割り当てられたマイクロタスクを実行するためにシステム108を使ってもよい。プロバイダー114は、種々の地理的位置に位置していてもよい。自動化システムまたは機械110はコンピュータ・システムおよびアプリケーションを含んでいてもよい。図1には分配システム106は一つしか描かれていないが、MMS 104は複数の分配システムと協働するよう構成されていてもよい。

#### 【0020】

MMS 104は、タスクについての一組のマイクロタスクに対応するマイクロタスク作業生成物を受け取るよう構成される。マイクロタスク作業生成物は、分配システム106から、あるいは一または複数の作業員から直接、受領されうる。MMS 104は、受領されたマイクロタスク生成物に基づいて、タスク要求において要求されていたタスクについての最終生成物(最終出力)を構築するよう構成される。ある実施形態では、MMS 104はマイクロタスク生成物を総合してそのタスクについての最終出力生成物を生成するよう構成される。タスクについての作業生成物は次いでタスク要求者に提供されうる。さらなる詳細は下記で与える。

#### 【0021】

先に論じたように、実行するには、あるいはより効率的もしくはより正確に実行するには、人間の知性を使って人間によるしかないタスクがある。結果として、これらのタスクは、人間の処理または人間と機械の処理の組み合わせを使って実行されるのが最善である。これらのタスクは典型的には、タスクについて提供される入力(入力情報)の解析、要約または処理に関わる。たとえば、文書の画像を与えられて、タスクは文書画像の記号表現を含む出力を生成することに関わるものでありうる。記号表現(symbolic representation)は、文書画像から抽出された単語および文書ページに描かれている何らかのオブジェクトの名称を含んでいてもよい。別の例としては、音声情報が与えられたとき、タスクはその音声情報の文字起こしを生成することに関わるものでありうる。さらにもう一つの例としては、入力情報は構造化されていない情報(たとえばテキスト・ファイル)を含んでいてもよく、タスクはその情報を構造化されたフォーマットにすること(たとえば情報をスプレッドシートに入力すること)であってもよい。

#### 【0022】

実行されるべき何らかのタスクについて、コスト、人間または機械作業員によってタスクを実行させることに関わるリスク、タスクの作業生成物の期待される品質などといったタスクに関連するいくつかの因子がある。たとえば、タスク入力に含まれる情報を人間または機械に与えることに関わるいくばくかのリスクがありうる。たとえば、入力情報が企業によって提供される名前、住所および電話番号の手書きのリストを含む場合、人間またはさらには外部の機械に管理される処理システムに入力情報を与えることは、企業のプライバシー・ポリシーに反する、あるいは該情報を扱う機関に対する信用の欠如につながることもありうる。このリスクは、タスクを実行する人間または機械が典型的にタスク要求者に知られない、自動化されたオンライン・アウトソーシング・シナリオでは特に高まる。

#### 【0023】

ある実施形態では、MMS 104は、タスクに関連しうるさまざまな因子を考慮に入れる仕方でタスクをアウトソーシングするよう構成される。たとえば、MMS 104は、タスクおよびその入力情報をアウトソーシングすることに関わるリスク(すなわち、アウトソーシングの結果として入力情報のプライバシーまたは秘匿性が危殆化されるリスク)を

10

20

30

40

50

減らすことに向けた、タスクのマクロタスクへの分割を含む、アウトソーシングを実行するよう構成される。MMS 104は、これを達成するためにさまざまな異なる技法を使ってもよい。たとえば、多くの場合において、MMS 104は、各作業者に与えられる入力情報を制御してもよい。これは、各作業者に暴露される入力情報の量を制限することによって、またさらには作業者に与えられる情報を修正することによってなされてもよい。

#### 【0024】

作業者に与えられる入力情報を制限および/または修正するために、さまざまなセグメント分割および/または組み合わせ技法が使用されうる。ある実施形態では、マクロタスクを実行する作業者に入力情報の部分集合しか与えられないようなセグメント分割技法が使われてもよい。作業者はそれでも、その作業者に与えられた入力情報の何らかの解析、要約または処理されたバージョンを提供することはできるが、入力情報全体はその作業者には利用可能でないため、タスクについて受領される入力情報の流布に関わる全体的なリスクは大幅に軽減される。たとえば、タスクについての入力情報が住所および電話番号に関わる名前の画像を含む場合、入力画像は三つの画像にセグメント分割されてもよい：第一の画像は名前だけを含み、第二の画像は電話番号だけを含み、第三の画像は住所だけを含む。次いで、各画像セグメントは別々の作業者に提供されてもよい。マクロタスクは、各画像セグメントに対応するテキスト情報を生成することである。このようにして、どの単一の作業者もすべての入力情報へのアクセスはもたない。これは、作業者によって知られる名前およびその関連付けられた住所および電話番号の流出に関わるリスクを減らす。

#### 【0025】

リスクをさらに軽減するため、マクロタスクのために作業者に与えられる入力情報は、組み合わせ技法を使うなど多様な仕方で修正できる。たとえば、上述したような入力情報が住所および電話番号が関連付けられた名前の画像を含む状況では、電話番号の画像だけがある作業者に与えられてもよい。秘匿情報が知られるリスクをさらに軽減するため、電話番号の画像は他の電話番号の画像、さらには偽の電話番号と組み合わせられてもよい。電話番号の修正された画像は次いで作業者に与えられてもよい。これは電話番号を作業者に開示することに関わるリスクを軽減する。作業者は、該作業者に与えられた入力情報がどのように修正されたかを全く知らないからである。それでも、作業者は、該作業者に割り当てられたマクロタスク（これはたとえば画像をテキスト情報に変換することであってもよい）を実行できる。

#### 【0026】

一般に、タスクに関わるリスクは次のように表現されうる。Iをタスクについての完全な入力情報とする。Oを入力Iに対してタスクを実行することから生成される完全な出力（または作業生成物）とする。完全な出力Oは、入力情報全体を単一の作業者Wに与えることによって得ることができる。このことを $O=W(I)$ と書く。しかしながら、単一の作業者に入力情報全体を与えることに関わるリスクRがある。リスクは入力情報のプライバシー/内容が危殆化されることに関するものである。このリスクRは、作業者および与えられる入力情報に依存し、 $R(I,W)$ と表してもよい。 $R(I,W)$ は受け容れられないほど高いことがありうる。全体的なリスクを軽減するために、入力Iを修正することができる。たとえば、入力Iは部分集合 $I_1, I_2, \dots, I_n$ に分割され、ある部分集合とそれに関連するマクロタスクが作業者 $W_i$ に与えられることができる。各作業者 $(W_i)$ は、その作業者に割り当てられたマクロタスクを実行し、受領された入力 $(I_i)$ に基づいて出力 $(O_i)$ を生成することができる。ここで、 $O_i = W_i(I_i)$ である。最終出力Oは、作業者 $W_i$ からの出力 $O_i$ をまとめたバージョンである。したがって、 $O = \text{Assemble}(O_1, O_2, \dots, O_n)$ である。そのような技法に関するリスクは、入力の部分集合を各作業者に与えることに関するリスクの組み合わせである。作業者間での情報交換のリスクが十分小さい場合、全体的なリスクRはほぼ $R = R(I_i, W_i)$ である。ここで、 $R(I_i, W_i)$ は、作業者 $W_i$ に部分集合 $I_i$ を与えることに関わるリスクである。ある実施形態では、MMS 104によってなされる入力情報修正は、 $R(I_i, W_i)$ が $R(I, W)$ より小さくなるようなものである。

## 【 0 0 2 7 】

タスクに関連しうるもう一つの因子は、タスク実行から帰結する出力の期待される品質である。したがって、作業員Wに入力Iを与えることによって得ることのできる、 $Q(I, W)$ で表せるある品質Qがある。一般に、品質は、「所望される」または「正しい」出力への近さと考えることができる。もちろん、多くの場合において、「正しい」出力は未知であり、推定することしかできない。タスクのアウトソーシングが、前記リスクを軽減するのみならず、タスク出力の品質を最大化もするような仕方で行われることが望まれる。

## 【 0 0 2 8 】

タスクに関連しうるさらにもう一つの因子は、タスクを実行するためのコストである。作業員に入力に対して作業させて出力を生成させることに関わるコストがある。このコストは、 $C(W_i, I_i)$ と表してもよい。典型的には、機械実装されるプロセスによってタスクを実行するコストは、人間の作業員のコストより低いが、出力の品質も低くなることもある。

10

## 【 0 0 2 9 】

記号Wは作業員について使われ、作業員は自動化プロセス（機械）であっても人間の作業員であってもよい。さまざまな作業員におけるリスクの差もある。たとえば、自動化プロセスに関連するリスク（すなわち、入力情報のプライバシーまたは秘匿性が作業員にタスクまたはマイクロタスクをアウトソーシングする結果として危殆化されるリスク）は、人間作業員に関連するリスクよりもしばしば低い。これは部分的には、自動化プロセスは、そのタスクについてのもとの入力Iを与えるエンティティによってよりよく管理されるためである。リスクは作業員および与えられる入力の両者とともに変わるので、ある作業員からの出力を別の作業員への入力として与える、あるいは入力のより大きな部分を機械作業員に与えより小さな部分を人間の作業員に与えることがリスクを軽減することがありうる。たとえば、機械作業員に与えられる入力の部分集合しか人間の作業員には与えられなくてもよい。たとえば、自動化プロセスは、入力画像におけるすべての単語に対して作用して、その多くについての記号出力を生成するが、入力画像中の単語のいくつかを、たとえば筆跡が汚いために処理できないことがありうる。この場合、自動化プロセスによって困難だった入力画像の部分だけが人間の作業員に入力として渡されてもよい。

20

## 【 0 0 3 0 】

一般には特定の作業員に関するリスクは初期には未知であるが、リスクは、経験に基づいて推定することができ、作業員位置、作業員教育レベル、作業員収入、作業員評価レーティングおよび作業員の他の多くの特性に依存しうる。これらの特性のいくつかは、種々の状況における作業員との経験を収集することによって、MMS 104によって時間がたつとともに推定されうる。いくつかの特性は、他の手段によって推定されうる。場合によっては、特定のタスク分配システムが作業員に関連付けられた契約、非開示協定または束縛を提供した場合にはリスクはより低いことがありうる。よって、特定の入力を特定の作業員に与えることに関するリスク $R(I, W)$ は、複数の因子に基づいて推定されてもよい。

30

## 【 0 0 3 1 】

アウトソーシングの一環として作業員に与えられる入力情報は、入力情報に「ノイズ」を加えることによって修正されてもよい。たとえば、入力人間作業員に与えられる前に入力に「ノイズ」を加えることによって入力を修正するために自動化プロセスが使われてもよい。上述したように、本物の電話番号の画像が人間の作業員に与えられる前に、偽の電話番号の画像が追加されることができる。この場合、マイクロタスクに対応する出力は、 $O_1 = W_{\text{machine}}(I_1)$ および $I_2 = O_1$ および $O_2 = W_{\text{human}}(I_2)$ と表されうる。したがって、機械（machine）に与えられる第一のマイクロタスクは入力情報 $I_1$ にノイズを加えて出力 $O_1$ を生成する。第一のマイクロタスクの出力 $O_1$ は次いで、人間（human）の作業員に割り当てられた別のマイクロタスクの入力（ $I_2$ ）として与えられる。この処理シーケンスでは、 $R(I_2, W_{\text{human}}) < R(I_1, W_{\text{human}})$ であり、 $R(I_1, W_{\text{machine}})$ は自動化プロセスであるため非常に小さいので、リスクが軽減される。

40

## 【 0 0 3 2 】

50

リスクを軽減するために加えることのできるもう一つの型のノイズは、作業者に与えられる（入力される）画像に対するぼかしまたはその他の歪みである。たとえば、入力画像は人間および認識されるべきオブジェクトを含んでいてもよい。画像がそのまま与えられれば、人間が誰であるかは認識されうる。しかしながら、何らかのぼかしが適切な位置において入力画像に適用されれば、作業者はその人間を同定することはできないが、それによって画像中のオブジェクト、たとえば自動車、空、建物などを同定することはできるとありうる。この場合、 $R(I,W) > R(\text{Distorted}(I),W)$ である。この場合、同定タスクの品質は、画像を歪めることによって損なわれることがありうる。この場合、 $Q(I,W) > Q(\text{Distorted}(I),W)$ である。よって、タスク要求者の期待に依存して、リスクと品質のパラメータ間でバランスを取ってもよい。

10

## 【0033】

「ノイズ」は、品質管理目的のために入力に加えられてもよい。たとえば、自動化プロセスによって注入される「ノイズ」は、所望される出力が既知である入力情報を含んでいてもよい。たとえば、記号的電話番号が既知である電話番号画像である。これは、出力の品質の評価を許容する。人間の作業者が既知部分について正しい答えを含まない出力を提供する場合、MMS 104はそのタスクについてのその作業者によってなされた作業に対して、低い品質推定を割り当てることができる。低い品質推定が要求される品質を下回る場合、タスクは別の作業者に割り当てられてもよく、該別の作業者が品質要求を満たすかどうかを見るために再び新しいジョブの品質が推定されることができ。

## 【0034】

20

作業者がすでに完遂した以前のタスクに関係した情報を、その同じ作業者に与えることに関わるリスクもある。これは、その作業者が情報を蓄積して潜在的にそれを流出させることを可能にするからである。たとえば、同じ作業者が社会保障カードおよび同じ人物からの申し込み書式の画像を与えられる場合、その作業者がその申込者についてもつ情報は多過ぎることになりうる。よって、一般に、 $R(I_1+I_2,W_0) > R(I_1,W_1)+R(I_2,W_2)$ である。

## 【0035】

先述したように、リスクRを最小にしつつ、最終出力の品質を最大にすることも望ましい。コスト制約のため、場合によっては両方を同時に行うことは可能でないこともあり、よって、何らかの受け容れ可能な出力品質および何らかの受け容れ可能なリスクをもつ動作点〔オペレーティング・ポイント〕を選んでもよい。さらに、品質およびリスクは式で記述されうるものの、精密なリスクおよび品質は各入力に依存し、よってリスクおよび品質は推定することしかできないことがありうる。よって、入力の分割の影響が精密にはわからないときに、これらの目標を達成しようとして使われうる一組の指針または規則が使用されてもよい。たとえば、より高いプライバシー（すなわち、リスクを下げるべき）が所望される場合、タスクはより多くのマイクロタスクに分割されてもよい。各マイクロタスクは、その入力として、入力全体の部分集合をもつ。したがって、ある実施形態では、所望されるプライバシー・レベルが高い（すなわち、許容できるリスクが低い）ほど、タスクについてのマイクロタスクの数は大きくなり、これはそのタスクについて与えられる入力情報のより大規模な分割を含意する。たとえば、文書全体をマクロタスクについての入力として一作業者に送ることにおいて推定される一つのリスク・レベル、その文書を二つの半分に分割して一方の半分を位置の作業者に、他方の半分を別の作業者に送ることに関わるより低いリスク・レベル、入力文書を三つ以上の部分に分割して（たとえば文書を一行ごとに分割して）各部分（たとえば各行）を別々の作業者に送ることに関する最小リスク・レベル、などがあってもよい。このように、受け容れ可能なリスクのレベルは、タスクのマイクロタスクへの分割がどのように実行されるべきかを決定するための因子として使われてもよい。

30

40

## 【0036】

多くの場合、複数の作業者の間での入力の分割はより低い品質につながる。たとえば、 $Q(I_i,W_i)$ は $Q(I,W)$ より低い。しかしながら、場合によっては、分割は品質を改善する（すなわち、 $Q(I_i,W_i) > Q(I,W)$ ）。たとえば、一作業者が電話番号を認識することにとり

50

わけ熟達しており、別の作業者は特定言語の文字を認識することにとりわけ熟達している場合、マイクロタスクが、該マイクロタスクを実行することに熟達している作業者に割り当てられるよう、タスクを分割すれば、リスクを下げつつ品質を上げることができる。

【 0 0 3 7 】

品質を上げるために複数の動作を使用できる。たとえば、複数の作業者が同じデータに対して同じマイクロタスクを実行するよう依頼され、その作業生成物が比較されて品質レベルを決定してもよい。別の状況では、ある作業者によってマイクロ作業実行の結果として生成された作業生成物が、別の作業者によって（別のマイクロタスクの一環として）検査されてもよい。追加的な作業者による追加的なマイクロタスクを使うことは、情報の分配を増すことになり、潜在的にはタスク完了に関する全体的なリスクを高めることになりうる。

10

【 0 0 3 8 】

入力の部分集合に対する作業の品質を改善するために人間の作業者が使用されることができ。たとえば、人間は、自動化プロセスの出力または作業生成物を与えられて、それをチェックするよう依頼されてもよい。まず、自動化プロセスが $O_1 = W_{\text{machine}}(I_1)$ を行い、次いで人間は $O_2 = W_{\text{human}}(I_1, O_1)$ を生成する。ここで、第二の出力は、自動化プロセスの訂正された形である。自動化プロセス（機械）によって実行されるタスクまたはマイクロタスクは、自動化プロセスの以前の動作に基づく何らかの期待される品質 $Q_{\text{machine}}$ をもつことができる。あるいは場合によっては、自動化タスクは入力に基づく品質を自己報告することができる。たとえば、文字を認識する自動化タスクは「c」である可能性が99%で「e」である可能性が1%と報告してもよく、これは中程度の品質と考えられてもよい。しかしながら、自動化タスクが「c」である可能性が90%で「e」である可能性が10%と報告する場合、これは低品質と考えられ、おそらくは追加的な処理または訂正を必要とする。人間は訂正しか行わないので、全体的なタスクに関するコストは、もとの入力に対して人間に作業させるよりも低い。よって、 $C(W_{\text{human}}, I_1, O_1) + C(W_{\text{machine}}, I_1) < C(W_{\text{human}}, I_1)$ である。

20

【 0 0 3 9 】

ある実施形態では、タスクまたはマイクロタスクの品質 $Q$ は自動化技法を使って測定されてもよい。一つのそのような自動化技法は、作業者 $W_i$ によって提出される単語数 $N_i$ および対応するマイクロタスク（たとえばある文書画像）において自動単語境界検出アルゴリズムによって検出される単語数 $M$ の関数であることができる。この場合、品質 $Q$ は $Q = f(M, N_i)$ として測定される。別の関数は $M$ と $N$ の比 $Q = M/N$ を使ってよく、別の関数は $Q = M - N/M$ であってもよい。ある実施形態では、二以上の作業者が特定のジョブについての結果を提出する場合、自動化品質検出は、すべての出力に依存する関数 $Q = f(N_1, \dots, N_i)$ であることができる。ある場合、 $f$ は、 $1 / \text{累積編集距離}(N_1 \dots N_i)$ であってもよい。もう一つの場合には、 $Q = (\min(N_1 \dots N_i) / \max(N_1 \dots N_i))$ である。

30

【 0 0 4 0 】

ある実施形態では、入力および出力が画像である場合（または画像に変換できる場合）、ジョブ品質 $Q$ は、入力画像の規格化されたグレーレベルまたはカラー・ヒストグラム $H_i$ を、タスク出力の規格化されたグレーレベルまたはカラー・ヒストグラム $H_o$ と比較することによって測定できる。この場合、 $Q = f(H_i, H_o)$ である。

40

【 0 0 4 1 】

同じ作業者が、すぐ連続して、あるいはある時間期間の間に、いくつかの同様の型のタスクを実行する場合、その作業生成物はその作業者によって生成される品質を推定するために使われることができる。最も単純なレベルでは、作業者品質は、タスクの以前の受入率（acceptance rate）に基づいていてもよい。作業者品質のより正確な推定は、先に論じたように、既知の入力に対する実績を考慮に入れてもよい。十分なデータがあれば、作業者品質推定は、作業者疲労（多くの連続したタスクの後の品質推定を下げる）または作業者が該作業者の時間帯〔タイムゾーン〕においてタスクを実行している時刻を考慮に入れることができる。特定の作業者についてその作業者の品質を推定するには不十分なデー

50

たしかない場合には、おおまかな初期推定は、他のタスクに対する実績、あるいはより多くのデータがある他の作業員との何らかの類似性、たとえば地理的位置、言語スキルに基づいていてもよい。タスクのさまざまな型についての品質推定を確立するために初期タスクを作業員に割り当てることも可能である。

#### 【0042】

ある実施形態では、システム100のMMS 104は上で論じた原理を具現する。さまざまな因子を考慮に入れつつタスクをいかにしてマイクロタスクに分割すべきかを制御する規則がMMS 104のために構成されてもよい。該因子は、そのタスクについての受け入れ可能なリスク閾値、そのタスクについての所望される品質レベル、そのタスクについての受け入れ可能なコスト閾値、マイクロタスクが人間または自動化された作業員によって実行されるべきかどうか、および上で論じた他の因子といったものである。システム100によって実行される処理に関する詳細は後述する。

10

#### 【0043】

図1に描かれるシステム100を参照するに、タスクまたは作業要求者は、システム102を使って、タスク要求112をMMS 104に通信してもよい。タスク要求者は人間、機械、ソフトウェア・アプリケーション/プロセスなどであってもよい。ある実施形態では、タスク要求者はタスク要求を構成するためにMMS 104と直接対話してもよい。MMS 104は複数のタスク要求者から複数のタスク要求を受け取ってもよい。このようにして、MMS 104はいつでも、複数の要求者にサービスし、複数のタスク要求を処理しうる。

20

#### 【0044】

MMS 104によって受領されるタスク要求112は、実行されるべきタスクを同定するタスク記述を含んでもよい。タスクは、人間知性タスク(HIT)または他のタスクであってもよい。タスク要求は、実行されるべき一つまたは複数のタスクを指定してもよい。要求されるHITの例としては、これに制約されるわけではないが、次のようなものが含まれる。

- ・手書き文字またはテキストを画像からタイプ・テキストに変換する(たとえば、一つまたは複数の名刺から連絡先情報をタイプする、顧客が記入した書式データをエクセル・スプレッドシートにタイプする、文書修正をタイプする、名刺からの情報をデータベースに格納されている連絡先情報にタイプする)。
- ・グラフィック(たとえば手書きのグラフィック、ロゴ)をコンピュータ描画に変換する(たとえば、画像からのグラフィックをVISIO描画データに変換する、ホワイトボード画像をパワーポイント・スライドに変換する)。
- ・オブジェクト、画像、文書に対して、メタデータによりタグ付けする/記述する(たとえば、写真に写っている人の名前を入力する)。
- ・オブジェクト、画像、文書を分類する(たとえば、文書が請求書か納税用紙か分類する)。
- ・デジタル貯蔵所においてオブジェクト、画像、文書を見つける(たとえば、文書Aのすべてのバージョンを見つける、人物Aについてのリンクトイン(Linked-in)のURLを見つける)。
- ・オブジェクト、画像、文書の間の関係を定義する(たとえば請求書Aが書類Bに関係する)。

30

40

#### 【0045】

MMS 104によって受領されるタスク要求112はまた、要求されたタスクを実行するために使われるべき情報、すなわちそのタスクについての入力情報をも含んでもよい。入力情報は、実行されるべきタスクに依存してもよく、一つまたは複数の型の情報を含んでもよい。一つまたは複数の型の情報には、これに限られないが、テキスト情報、画像情報、オーディオ情報、ビデオ情報、グラフィック、手書き情報および他の型の情報ならびにそれらの組み合わせが含まれる。タスクについての入力情報は、さまざまな異なる形において与えられてもよい。ある実施形態では、入力情報は、一つまたは複数の文

50

書の形で与えられ、各文書が一つまたは複数の型の情報を含んでいてもよい。入力文書は、テキスト・ファイル、スキャナによって生成されたファイル、ワードプロセッシング・プログラムによって生成されたファイル、画像または写真、オーディオ・ファイル、ビデオ・ファイルなどであることができる。たとえば、入力文書は名刺、レシート、手紙のメモ、ラベル、標識、請求書、写真、書式または図、新聞記事、小切手、オブジェクトなどの画像であってもよい。

【0046】

上に示したように、タスクについて与えられる入力情報は典型的には実行されるべきタスクに依存する。たとえば、実行されるべきタスクが音声文字起こしである場合、入力情報は、文字起こしされるべき一つまたは複数の音声ファイルを含んでいてもよい。もう一つの例として、タスクが第一の言語から第二の言語に翻訳することである場合、入力情報は、翻訳されるべき第一の言語における一つまたは複数の文書を含んでいてもよい。もう一つの例としては、タスクが画像中のオブジェクトを識別する/タグ付けすることである場合、入力文書は一つまたは複数の画像であってもよい。したがって、タスクについて与えられる入力情報の内容は、実行されるべきタスク（または入力情報が複数のタスクについて使用されるのであれば複数のタスク）の型に依存しうる。

【0047】

図1を参照するに、タスク記述および入力情報に加えて、MMS 104によって受領されるタスク要求112は、実行されるべきタスクに関係する他の情報をも任意的に含んでいてもよい。たとえば、タスク要求は、タスクを実行するために使用されるべき基準を指定してもよい。該基準は、タスクを実行するための所望される価格/コストに関する情報、タスクに関連するリスクの所望されるまたは受け容れ可能なレベル、タスクを完了するための時間枠、タスク出力についての所望される品質、タスクを実行する作業者の所望される位置、使われる作業者のタイプ（たとえば人間か機械か）、分配制約条件およびその他の情報といったものである。すると、この追加情報は、マイクロタスクを生成する、入力情報をセグメント分割する、マイクロタスクを分配するなどのためにMMS 104によって使用される。

【0048】

ある実施形態では、タスク要求112は、プライバシーが保護されるべき入力情報の一つまたは複数の部分を同定する情報をも含んでいてもよい。これは、タスク要求者が、要求者にとって重要であり、プライバシーが保護されるべきである入力情報の部分を具体的に特定することを可能にする。たとえば、スキャンされた画像（たとえば名刺の画像）の内容に対応するテキストを生成することに関わるタスクについては、タスク要求者は人物の名前、人物の雇用者および雇用者の住所のプライバシーが保護されるべきであることを指定してもよい。タスク要求112はまた、供給されるデータに関して取ることのできる受け容れ可能なリスクの全体的なレベルをも含んでいてもよい。この情報は、MMS 104によって、そのタスクについて指定されたさまざまな因子（たとえば、リスク、品質、コストなど）を満足しつつ、タスクをどのようにしてマイクロタスクに分割するかを決定するために使われてもよい。

【0049】

MMS 104は、MMS 104によって実行されるさまざまな機能を容易にするいくつかのサブシステムを含んでいてもよい。図1に描かれる実施形態では、MMS 104は、ユーザー・インターフェース・サブシステム116、内容解析サブシステム117、セグメント分割器サブシステム118、組み合わせ器サブシステム120、マイクロタスク生成器サブシステム122、値付け〔費用評価〕サブシステム124、前置処理器サブシステム126およびタスク生成物管理サブシステム（TPMS: task product management subsystem）128を有する。MMS 104のこれらさまざまなサブシステムはハードウェア、ソフトウェア（たとえばMMS 104のプロセッサによって実行されるコード、プログラム、命令）またはそれらの組み合わせにおいて実装されてもよい。ソフトウェアはコンピュータ可読記憶媒体に記憶されていてもよい。各サブシステムによって実行される処理

10

20

30

40

50

は、のちにさらに詳細に記述する。

【 0 0 5 0 】

MMS 104によって実行される処理を制御する一組の規則が、MMS 104について構成されてもよい。これらの規則は、リスク・レベル、品質、コストなどに関係するようなタスクについて指定されるさまざまな因子または制約条件が満たされるようにタスク・アウトソーシングを実行するために、MMS 104によって実行されるさまざまな処理機能を制御する。たとえば、リスクに関し、これらの規則は、タスクについて指定されるリスク・レベルが満足されるように特定のタスクをどのようにマイクロタスクに細分するかを制御する。もう一つの例として、品質に関し、これらの規則は、タスク実行から生成される出力の全体的な品質が指定された品質閾値を満たすよう、どのように特定のタスクをマイクロタスクに細分するかを制御する。リスクを品質とバランスする必要がある状況では、これらの規則は、品質を最大にしつつリスク・レベルが満足されるよう、あるタスクについてのマイクロタスクを決定するために使用されてもよい。図1に描かれる実施形態については、一組の規則は、タスク規則130、セグメント分割規則132、組み合わせ規則138、マイクロタスク規則144、分配規則148および価格付けサブシステム124によって使われる規則を含む。

10

【 0 0 5 1 】

タスクに影響し、該タスクのマイクロタスクへの分割に影響する因子が指定されるさまざまな方法がある。上記のように、タスク要求者は、タスク要求においてこれらの因子の一つまたは複数を指定してもよい。あるいはまた、デフォルト因子がMMS 104について構成されてもよい。MMS 104は、実行されるべきタスクの性質に基づいて（たとえば、タスク自身、入力情報の性質などに基づいて）、タスクについて使われるべき因子の組を決定するよう構成されてもよい。

20

【 0 0 5 2 】

MMS 104は任意的に、MMS 104に情報を提供し、MMS 104が情報を出力する（たとえばタスク要求者システム102またはタスク要求者に）ためのインターフェースを提供するよう構成されたユーザー・インターフェース・サブシステム116を有してもよい。ある実施形態では、ユーザー・インターフェース・サブシステム116は、タスク要求者のようなユーザーがMMS 104と対話できるようにするグラフィカル・ユーザー・インターフェース（GUI）の組を提供してもよい。たとえば、タスク要求者がタスク要求を構成設定できるようにする、タスク要求に関連した情報を与える、タスクについての最終作業生成物を閲覧するなどできるようにするGUIが設けられてもよい。タスク要求者がMMS 104によって実行される処理を構成設定できるようにするGUIが設けられてもよい。たとえば、タスク要求者は、ユーザー・インターフェース116を使って、MMS 104の一つまたは複数のコンポーネントによって実行される処理に影響する規則または基準を構成設定してもよい。

30

【 0 0 5 3 】

ユーザー・インターフェース・サブシステム116はまた、MMS 104の動作を制御するためにMMS 104のユーザーが使用しうる一組のアプリケーション・プログラミング・インターフェース（API）をも提供してもよい。たとえば、タスク要求者が、実行されるべきタスク、そのタスクについての入力情報およびそのタスクがどのように実行されるべきかに関する他の基準を指定できるようにするAPIが提供されてもよい。ある実施形態では、ユーザー・インターフェース・サブシステム116は、タスク要求112を受領し、タスク要求をさらなる処理のために内容解析サブシステム117に転送するよう構成される。

40

【 0 0 5 4 】

内容解析サブシステム117は、タスク要求および関連する情報を解析するよう構成される。解析の一環として、内容解析サブシステム117は、実行されるべきタスクの型を認識する、タスクについて受領された入力の型を判別する、実行されるべきタスクに対して課された制約条件があればこれを判別するなどするよう構成されてもよい。解析を実行

50



することから内容解析サブシステム 117 によって探り出される情報は、次いで、MMS 104 の他のサブシステムによって使用されてもよい。たとえば、内容解析サブシステム 117 によって実行される解析は、MMS 104 の他のサブシステムによって、そのタスクについてのマイクロタスクを決定することに関係する処理について使用されるべき規則を選択するために使用されてもよい。たとえば、内容解析サブシステム 117 によって実行される解析は、そのタスクについてセグメント分割器 118 によって使用されるべきタスク規則および/またはセグメント分割規則を決定するために、セグメント分割器 118 によって使用されてもよい。たとえば、内容解析サブシステム 117 は、タスクについての入力画像であることを判別し、次いで入力画像をホワイトボード画像、名刺、文書画像などとしてさらに分類するための一つまたは複数のアルゴリズムを実行してもよい。内容解析サブシステム 117 によって決定された分類情報は、次いで、入力画像をセグメント分割するための適切なセグメント分割規則を選択するために、セグメント分割器 118 によって使用されてもよい。

10

**【0055】**

状況によっては、タスク要求は実行されるべきタスクすら同定せず、入力情報だけを指定してもよい。内容解析サブシステム 117 は、入力情報を解析し、実行すべきタスクを自動的に決定してもよい。内容解析サブシステム 117 は、実行すべきタスクを自動的に決定するためにタスク規則 130 を使ってもよい。

**【0056】**

タスク要求は、タスクに関連付けられるべき一つまたは複数の因子または制約条件（たとえばリスク、品質など）を同定してもよい。内容解析サブシステム 117 は、これらの制約条件を認識し、その情報を MMS 104 の他のサブシステムに伝達するよう構成されていてもよい。ある実施形態では、内容解析サブシステム 117 は、解析に基づいて、そのタスクのために使われるべき一組の因子を決定するよう構成されていてもよい。たとえば、リスク・レベルがタスク要求において指定されていない場合、内容解析サブシステム 117 によって実行される解析が、そのタスクに関連付けられるべきリスク・レベルまたは閾値を決定するために使われてもよい。

20

**【0057】**

上に示したように、内容解析サブシステム 117 は、タスク規則 130 を使って、実行すべきタスクを自動的に決定してもよい。ある実施形態では、タスク規則は、条件と、該条件が満たされるとき、あるいは該条件が満たされないときに実行されるべきタスクとを特定してもよい。タスク規則のための条件は、タスク要求者の素性〔識別情報〕、入力情報に含まれる情報の型および他の基準ならびにそれらの組み合わせといった一つまたは複数の基準に基づいていてもよい。タスク規則の例：

30

(1) 入力情報 = 音声情報のみであれば、タスク = 音声情報を文字起こしする；

(2) 入力情報のソース = ユーザー A かつ 入力情報 = 一つまたは複数のスキャン画像であれば、タスク = 各入力画像のテキスト内容をタイプ・テキストに変換し、各画像のグラフィック内容をコンピュータ描画に変換する。

タスク規則 130 はユーザー構成設定可能である。たとえば、ユーザー・インターフェース・サブシステム 116 によって提供される API または GUI が、タスク要求者によって、該要求者のためのタスク規則をカスタマイズするために使用されてもよい。

40

**【0058】**

MMS 104 の機能の一つは、タスクについて受領される入力情報の秘匿性およびプライバシーを保護しつつ、タスクの実行をアウトソーシングすることである。上で論じたように、これを行う一つの方法は、タスクをマイクロタスクに分割して、各マイクロタスクには入力情報の部分集合が関連付けられるようにし、次いでマイクロタスクを複数の作業者にアウトソーシングすることである。タスクをマイクロタスクに分解することによって、入力情報は部分集合にセグメント分割され、各部分集合が、ある作業者に割り当てられるマイクロタスクについての入力となる。ある実施形態では、入力情報の、マイクロタスクに関連付けられるべき部分集合へのセグメント分割は、セグメント分割器サブシステム

50

118によって実行される。セグメント分割器サブシステム118は、タスクについて受領される入力情報を一つまたは複数のセグメント136にセグメント分割するよう構成される。各セグメントは、入力情報の一部分を含む。

【0059】

セグメント分割器サブシステム118は、実行されるべきタスクおよびそのタスクに関連付けられた制約条件に基づいて、入力情報をセグメント分割するために、種々の型のセグメント分割技法を使ってもよい。使用されうるセグメント分割技法の例は、これに制約されるものではないが、さまざまな内容ベースのセグメント分割技法、時間的データ（たとえばビデオまたはオーディオ）については時間的セグメント分割技法その他を含む。内容ベースのセグメント分割技法の例は、これに制約されるものではないが、単語境界セグメント分割、画像/グラフィック・ベースのセグメント分割、キャラクタ・セグメント分割、キャラクタ行セグメント分割、領域セグメント分割、フェース（face）・セグメント分割、描画領域セグメント分割、手書きセグメント分割、オブジェクト・セグメント分割、署名セグメント分割などを含む。入力情報がオーディオ・クリップまたはビデオ・クリップのような時間的情報を含む場合、セグメント分割は、時間次元に沿って実行されうる。たとえば、オーディオおよびビデオ・クリップは、固定時間間隔に基づいてセグメント分割されてもよい。内容ベースのセグメント分割も時間的入力情報に対して実行されてもよい。セグメント分割技法は、全自動化されていてもよいし、あるいは自動化および手動の入力セグメント分割技法の組み合わせを含んでいてもよい。さらに、入力情報が選択されたセグメント分割技法を使ってセグメント分割されうる種々の方法がある。

【0060】

ある実施形態では、セグメント分割器サブシステム118は、セグメント分割規則132を使って、タスクについての入力情報をセグメント分割するために使われるべき一つまたは複数のセグメント分割技法を決定し、また、選択された一つまたは複数のセグメント分割技法を使って入力情報がセグメント分割される仕方を決定する。セグメント分割規則は、条件と、使用されるべき一つまたは複数のセグメント分割技法と、前記条件が満たされたとき（あるいは前記条件が満たされないときに）選択された一つまたは複数のセグメント分割技法を使って入力情報がセグメント分割される仕方とを特定してもよい。セグメント分割規則についての条件は、実行されるべきタスク、タスク要求者の素性、入力情報に含まれる情報の型（たとえばオーディオ情報、ビデオ情報、画像など）および他の基準ならびにそれらの組み合わせといった、一つまたは複数の基準に基づいていてもよい。セグメント分割規則の例：

（1）入力情報 = オーディオ情報のみであれば、セグメント分割技法 = 時間的セグメント分割技法A；

（2）タスク = テキストをタイプ・テキストに変換、かつ入力情報 = 画像であれば、セグメント分割技法 = 単語境界セグメント分割。

【0061】

ある実施形態では、タスクに関連付けられた受け容れ可能なリスク・レベルが、タスクがどのようにマイクロタスクに分解されるべきか、および、入力情報がどのように部分集合にセグメント分割されるべきかを制御してもよい。たとえば、入力情報がセグメント分割されるセグメントの数は、ジョブに関連付けられたリスク閾値に反比例してもよい。低いリスク閾値だと、入力情報はX個のセグメントにセグメント分割され、中くらいのリスク閾値だと、入力情報はY個のセグメントにセグメント分割され、高いリスク閾値だと、入力情報はZ個のセグメントにセグメント分割され、 $X > Y > Z$ である。さまざまなリスク閾値と対応するセグメント数との間のこの相関は、セグメント分割規則132においてエンコードされていてもよく、セグメント分割を実行するためにセグメント分割器118によって使用されてもよい。

【0062】

ある種の実施形態では、タスクに関連付けられる所望される品質レベルは、そのタスクについて受領された入力があるようにして、それぞれがマイクロタスクへの入力として与

10

20

30

40

50

えられる部分集合にセグメント分割されるかを決定するためにも使われてもよい。さまざまな品質閾値と、対応するセグメント数との間のこの相関は、セグメント分割規則 1 3 2 においてエンコードされていてもよく、セグメント分割を実行するためにセグメント分割器 1 1 8 によって使用されてもよい。

#### 【 0 0 6 3 】

セグメント分割規則 1 3 2 はユーザー構成設定可能である。たとえば、ユーザー・インターフェース・サブシステム 1 1 6 によって提供されるAPIまたはGUIは、タスク要求者がセグメント分割規則を、該要求者のニーズに適合するようカスタマイズできるようにしてもよい。ある実施形態では、異なるタスク要求者について異なるセグメント分割規則が構成設定されてもよい。特定のタスクについて、セグメント分割器サブシステム 1 1 8 は、そのタスクについて使われるべき一つまたは複数のセグメント分割規則を決定し、選択されたセグメント分割規則に基づいて使用されるべき前記一つまたは複数のセグメント分割技法および選択された技法を使って入力情報がセグメント分割される仕方を決定するよう構成される。セグメント分割器サブシステム 1 1 8 は、次いで、選択されたセグメント分割規則によって指定される仕方で、選択された技法を使って、入力情報をセグメント分割するよう構成される。セグメント分割を実行することから生成されるセグメント 1 3 6 は次いで、セグメント分割器サブシステム 1 1 8 によって、さらなる処理のために組み合わせ器サブシステム 1 2 0 に提供されてもよい。

#### 【 0 0 6 4 】

図 8 A および図 8 B は、本発明のある実施形態のさまざまな側面を記述するために、本願を通じて使用される例を描いている。しかしながら、この例は、請求項に記載される本発明の実施形態の範囲を限定することを意図したものではない。図 8 A に描かれる例では、タスクは画像のテキスト内容をタイプ・テキスト〔テキスト・データ〕に変換し、画像のグラフィック内容があればそれをコンピュータ描画（たとえばVISIO描画データ）に変換することであると指定するタスク記述 8 0 0 を含むタスク要求が受領されてもよい。そのようなタスクについて受領される入力情報 8 0 2 は、一つまたは複数の名刺のラスタ化された画像を含んでいてもよい。記述 8 0 0 は、供給される入力情報に関して取ることのできるリスクのレベルおよび/または所望される品質のレベルといった、そのタスクに関連付けられる因子または制約条件を指定する情報を含んでいてもよい。図 8 A の例では、入力情報は、二つの名刺の二つのスキャンされた画像 8 0 4 および 8 0 6 を含む。ある実施形態では、スキャンされた画像は、タスク要求者システムからMSS 1 0 4 によって受領されるタスク要求に含まれていてもよい。別の実施形態では、MMS 1 0 4 は画像を生成するための機構を提供してもよい。たとえば、MMS 1 0 4 は、名刺を入力として取り、名刺をスキャンして画像 8 0 4 および 8 0 6 を生成し、次いでそれらの画像をタスクを実行するために利用可能にするスキャナを有していてもよい。

#### 【 0 0 6 5 】

図 8 A の例では、各入力画像は単語境界セグメント分割技法およびグラフィック・セグメント分割技法を使ってセグメント分割される。グラフィック・セグメント分割技法は、各入力画像中のグラフィックの位置を判別し、グラフィック部分だけを含むセグメントを生成する。単語境界セグメント分割技法は、各入力画像中の単語境界を判別し、単語境界に対応するセグメントを生成する。ある実施形態では、単語セグメント分割技法は、各入力文書画像中に、セグメント分割された単語領域を含む一組の長方形を識別する。セグメント分割された各単語領域は次いで、入力画像から抽出され、新しい画像として保存される。これらの新しい画像は入力画像についてのセグメントを表している。図 8 A に示されるように、画像 8 0 4 は、画像セグメント 8 1 0 と、単語境界に基づく12個の単語セグメント 8 1 2 とを含む13個のセグメント 8 0 8 に分割される。画像 8 0 6 は、画像セグメント 8 1 6 と、単語境界に基づく11個の単語セグメント 8 1 8 を含む12個のセグメント 8 1 4 に分割される。セグメント 8 0 8 および 8 1 4 は、次いで、さらなる処理のために組み合わせサブシステム 1 2 0 に提供されてもよい。ある実施形態では、セグメント分割を実行するために、非特許文献 1 に記載される技法が使用されてもよい。非特許文献の内容全

10

20

30

40

50

体はここにあらゆる目的について参照によって援用する。

【0066】

セグメント分割器サブシステム118が入力情報をセグメント分割する仕方は異なるタスクについては異なることがあるので、セグメント分割器サブシステム118は、各タスクについて、そのタスクについて入力情報がセグメント分割された特定の仕方を同定するセグメント分割情報134を記憶する。ある実施形態では、タスクについて記憶されるセグメント分割情報134は：タスクを同定する情報、タスクの要求者を同定する情報、タスクについて受領された入力情報を同定する情報、生成されるセグメント数を含め入力情報がどのようにセグメント分割されたかを同定する情報およびセグメントが生成された仕方（たとえば使用されたセグメント分割技法（単数または複数））、もとの入力中でのセグメントの位置および他の情報を含んでもよい。タスクについて受領された入力情報は複数の入力文書を含むことができるので、セグメント分割情報134は、それらの文書を同定する情報および、各文書について、その文書の内容がどのようにセグメント分割されたかを同定する情報を含んでもよい。このようにして、セグメントが与えられたとき、セグメント分割情報134は、そのセグメントに対応する入力文書と、該入力文書が入力として受領された目的のタスクとを判別するために使用されることができる。後述するように、セグメント分割情報134は、タスク生成物管理サブシステム（TPMS）128によって、タスクについての最終作業生成物を構築するために使用される。

10

【0067】

セグメント分割システムは、タスクをどのようにセグメント分割するかを決定するために、所望される入力リスクおよび品質を使ってもよい。たとえば、リスク要件がなければ、すべてのテキストは一つのセグメント内に留まってもよい。中くらいのレベルのリスクが許容されるならば、テキストは種々のセグメントにセグメント分割されることができる。非常に低いリスクが許容されるなら、入力是个々の単語にまでセグメント分割されることができる（図8Aに示されるように）。きわめて低いリスクのためには、入力は文字にまでセグメント分割されることができる。セグメント分割の選択は、所望される品質レベルを考慮に入れてもよい。一つの場合には、すべてのテキストを単一のセグメントに残すことは、高い品質につながる可能性がある。テキスト全体について同じフォントが使用されていることがあり、会社名のような情報が複数回、たとえば住所および電子メール・アドレスの両方に、現れることがあるからである。テキストを単語にセグメント分割することは、典型的には、文字に分割するよりは高い品質につながる。人間も機械も、単語のコンテキストにおいて文字を認識することの恩恵を受けるからである。

20

30

【0068】

タスクについての入力情報がセグメント分割され、セグメント分割情報が記憶されると、セグメント分割器サブシステム118がセグメント136を、さらなる処理のために組み合わせ器サブシステム120に提供してもよい。たとえば、セグメント分割器サブシステム118は図8Aに描かれるセグメント808および814を、さらなる処理のために組み合わせ器サブシステム120に提供してもよい。図8Aでは、単語画像セグメントは、境界で区切られたボックス内の単語（これは、単語画像セグメントに対応する図8Bにおける記号的な単語（symbolic word）とは区別する必要がある。記号的な単語は境界で区切られたボックスなしに示されている）によって表されている。

40

【0069】

組み合わせ器サブシステム120は、セグメント136から組み合わされたセグメント140を生成するよう構成される。ある実施形態では、組み合わせは、そのタスクについて受領された入力情報の内容のプライバシーを危殆化するリスクを軽減しようとするようになされる。セグメントが組み合わされる仕方は、実行されるべきタスクに関連付けられた受け入れ可能なリスクおよび/または品質のレベルといったさまざまな因子に依存してもよい。セグメント136の内容が、入力情報の内容のプライバシーを保護するよう（換言すれば、アウトソーシングの結果として入力情報のプライバシーまたは秘匿性が失われることに関連するリスクを軽減するよう）組み合わせ器サブシステム120によって組み

50

合わされうる種々の方法がある。ある実施形態では、この組み合わせは、入力情報中の情報間のコンテキスト関係をばかす方向でなされる（下記で例を与える）。別の実施形態では、入力情報の内容とともに「ノイズ」情報が組み合わせられてもよい。たとえば、偽の名前、電話番号などが、入力情報から抽出された内容に加えられてもよい。組み合わせ器は、種々の組み合わせ規則に関連付けられたリスクの推定を行って、それをそのタスクについて指定されているリスクのレベルと比較してもよい。

【0070】

ある実施形態では、組み合わせ器サブシステム120は、組み合わせ規則138を使って、どのように組み合わせが行われるかを決定する。組み合わせ規則は、条件と、該条件が満たされるときに（あるいは該条件が満たされないときに）使われるべき一つまたは複数の組み合わせ技法とを特定してもよい。組み合わせ規則についての条件は、一つまたは複数の基準に基づいていてもよい。下記の表Aは、組み合わせ器サブシステム120によって組み合わせが実行される仕方に影響しうる例示的な基準と、各基準の組み合わせ処理に対する影響を挙げている。

10

表A

基準 組み合わせ処理に対する影響

プライバシー考慮 組み合わせられた諸セグメントは、プライベートな情報が別個の組み合わせられたセグメントを横断してスクランブルされ、個々のどの組み合わせられたセグメントの内容も全体的な入力情報のプライバシーを危殆化させないように仕方で形成されうる。

20

たとえば、入力情報のある部分がプライベートであるとタグ付けされている場合（たとえば、先述したように、タスク要求者が入力情報のある部分をプライベートであると指定してもよい）、組み合わせ器サブシステム120は、組み合わせられたセグメントを生成するときにこれを考慮に入れてもよい。ある実施形態では、組み合わせられた諸セグメントは、プライベートとタグ付けされた入力情報内容が複数の組み合わせられたセグメントにまたがって分散され、単一の組み合わせられたセグメントの内容が知られても入力情報のプライバシーが危殆化されないように形成されうる。

【0071】

タスク要求または入力情報の源 タスク要求者は、該タスク要求者に由来するタスクについて使用されるべきカスタマイズされた組み合わせ技法を指定することが許容されてもよい。

30

【0072】

入力情報に含まれる内容の型 ある特定の組み合わせ技法がある種の情報については好適であるが、別の種類の情報については好適でないことがありうる。

【0073】

タスクに関連付けられた値付け情報 値付け情報は、生成されるべき組み合わせられたセグメントの数および組み合わせられたセグメントの内容に対して、タスクを完了するための全体的なコストがそのタスクについて指定された受け容れ可能な値付け限界内になるよう、影響しうる。

40

後述するように、各組み合わせられたセグメントについて一つまたは複数のマイクロタスクが決定される。次いで、各マイクロタスクを完了するための価格/費用が決定されてもよい。したがって、組み合わせられたセグメントの数はマイクロタスクの数に影響し、マイクロタスクの数はタスク完遂の全体的なコストに影響する。いくつかの事例では、組み合わせられたセグメントの数、よってマイクロタスクの数をある閾値を超えて増やすことは、そのタスクを実行するためのコストが受け容れ可能な値付けコストを超過することにつながる可能性がある。

【0074】

タスク完遂までの所望時間 タスク完遂までの時間は生成されるべき組み合わせられたセグメントの数および組み合わせられたセグメントの内容に影響しうる。組み合わせられた

50

セグメントの数はそのタスクについてのマイクロタスクの数に影響し、マイクロタスクの数はそれらのマイクロタスクを実行するために必要とされる時間に影響しうる。いくつかの事例では、多数の組み合わせられたセグメントを生成して各セグメントの内容を少なくすることは、全体的なタスクがより速く完了されることにつながりうる。だが、他の事例では、これは全体的なタスク完了に対して遅延を追加することがある。

**【 0 0 7 5 】**

作業者の利用可能性 上述したように、組み合わせられたセグメントの数はマイクロタスクの数に影響する。マイクロタスクは完了するために作業者にアウトソーシングされる。したがって、マイクロタスクの数はそれらのマイクロタスクを実行するために必要とされる作業者の数に影響する。したがって、組み合わせられたセグメントの数は作業者の利用可能性に基づいて決定されてもよい。

10

**【 0 0 7 6 】**

作業者のスキル・レベル 既知のスキル・セットをもつ作業者のために、組み合わせられたセグメントが生成されてもよい。それにより、該組み合わせられたセグメントに関連付けられるマイクロタスクがそれらの作業者によって実行されることができ。

**【 0 0 7 7 】**

人手 / 自動化 ある実施形態では、人間によって実行される必要があるタスクに関係した内容が一組の組み合わせられたセグメントにグループ化されてもよく、一方でコンピュータによって実行できるタスクが別個の組の組み合わせられたセグメントにグループ化されてもよい。このようにして、全体的なタスクは、人間の作業者および自動化されたコンピュータ技法のハイブリッドを使って実行できる。

20

**【 0 0 7 8 】**

タスクについての所望される精度 タスクについての所望される正確さの度合いは、生成される組み合わせられたセグメントの数および組み合わせられたセグメントの内容に影響しうる。

**【 0 0 7 9 】**

セグメントが組み合わせられる仕方は、タスクに関連付けられた受け入れ可能なリスク・レベルおよび / または品質レベルに依存してもよい。さまざまなリスクおよび / または品質閾値と使用されるべき組み合わせ技法との間のこの相関は、組み合わせ規則 1 3 8 においてエンコードされていてもよく、組み合わせを実行するために組み合わせ器 1 2 0 によって使用されてもよい。

30

**【 0 0 8 0 】**

図 8 A に描かれる例を参照するに、四つの組み合わせられたセグメント 8 2 0、8 2 2、8 2 4 および 8 2 6 がセグメント 8 0 8 および 8 1 4 の内容に基づいて生成されている。(グラフィックを含む) セグメント 8 1 0 および 8 1 6 の内容は単一の組み合わせられたセグメント 8 2 0 に組み合わせられている。組み合わせられたセグメント 8 2 2、8 2 4 および 8 2 6 は、入力画像から抽出された 8 1 2 および 8 1 8 における単語セグメントに基づいて生成されている。たとえば、セグメント 8 2 2、8 2 4 および 8 2 6 のそれぞれは、一つまたは複数のセグメント画像を含む新しい組み合わせられた画像(たとえば 8 2 2)を生成するよう、8 1 2 および 8 1 8 からのセグメント画像の一つまたは複数組み合わせるまたは融合することによって生成される。

40

**【 0 0 8 1 】**

組み合わせられた画像は、複数の異なる文書からのセグメント画像を含んでいてもよい。たとえば、組み合わせられたセグメント 8 2 2 は、文書 8 0 4 から抽出されたセグメント「Tom」〔トム〕を含み、文書 8 0 6 から抽出されたセグメント「Smith」〔スミス〕をも含む。同様に、組み合わせられたセグメント 8 2 0 は文書 8 0 4 からのセグメント 8 1 0 および文書 8 0 6 からのセグメント 8 1 6 を含む。同様に、組み合わせられたセグメント 8 2 4 および 8 2 6 も、文書 8 0 4 および 8 0 6 の両方からの内容を含む。代替的な実施形態(図 8 A には示さず)では、組み合わせられたセグメントは、意図的に導入されたノイズ内容をも含む。

50

## 【 0 0 8 2 】

図 8 A において見て取れるように、セグメントは、タスク要求者によってプライベートであるとタグ付けされた入力情報の部分（たとえば、人の名前、人の地位、人の雇用主および雇用主の住所）が三つの組み合わせられたセグメント 8 2 2、8 2 4 および 8 2 6 にまたがって拡散されるように、組み合わせられている。結果として、それぞれの組み合わせられたセグメントは、どの一つの入力文書からの情報全部よりも少ない部分を含んでいる。たとえば、入力文書 8 0 4 からの単語「Tom」が組み合わせられたセグメント 8 2 2 に含まれていたとしても、トムの名前「Jones」〔ジョーンズ〕は含まれない。組み合わせられたセグメント 8 2 2 の内容を見ている誰かにとって、名前 Tom Jones を判別することは可能ではなく、それにより名前のプライバシーが保持される。これは、初期タスクのそれぞれについて 10 の推定されるリスクを下げる。こうして、単一の名刺からのフルネームの流出のリスク 0.05 があった場合、ファーストネームと名字が分離されているときにシステムがフルネームを開示するリスクは 1000 分の 1、すなわち 0.00005 と推定されうる。これらの推定は経験的な結果に基づいていてもよい。

## 【 0 0 8 3 】

さらなるレベルのプライバシー保護が、入力文書から抽出された個々のセグメントが組み合わせられて組み合わせられたセグメントを形成する仕方によって可能にされる。たとえば、組み合わせられたセグメント 8 2 2 は文書 8 0 4 からのファーストネーム「Tom」と入力文書 8 0 6 からの名字「Smith」を含むセグメントを含む。セグメントのそのようなスクランブルは、さらなる層のぼかし（obfuscation）、そしてプライバシーが保持されるべき情報についての追加的な保護を加える。組み合わせられたセグメントの一つだけから実際の情報を見きわめることは、不可能ではないまでも、極めて困難だからである。

## 【 0 0 8 4 】

図 8 A に描かれた例から明らかなように、組み合わせられたセグメントは複数の入力文書から抽出されたセグメントを含むことができる。たとえば、組み合わせられたセグメント 8 2 2 は画像 8 0 4 からのセグメント「Tom」および画像 8 0 6 からのセグメント「Smith」を含む。したがって、二つ以上の異なる入力文書からの内容が単一の組み合わせられたセグメント中に組み合わせられてもよい。複数の入力文書にまたがる内容のこのようなスクランブルは、いずれか一つの文書からの内容のプライバシーが危殆化される可能性をさらに低下させる。

## 【 0 0 8 5 】

組み合わせ器サブシステム 1 2 0 は異なるタスクについて異なる組み合わせ技法を使ってもよいので、組み合わせ器サブシステム 1 2 0 は、各タスクについての、組み合わせられたセグメントを生成するためにそのタスクについてセグメントが組み合わせられた特定の仕方を同定する組み合わせ情報 1 4 2 を記憶する。ある実施形態では、タスクについて記憶されている組み合わせ情報 1 4 2 は、組み合わせられた各セグメントについて記憶されている以下の情報を含んでもよい：組み合わせられたセグメントを同定する情報、組み合わせられたセグメントと、該組み合わせられたセグメントに含められた諸セグメントとの間のマッピングを与える情報、組み合わせられたセグメント内の諸セグメントの位置およびその他の情報。したがって、組み合わせられたセグメントを与えられたとき、組み合わせ情報 1 4 2 を使って、該組み合わせられたセグメントに内容が含まれている一つまたは複数のセグメントを決定することができる。後述するように、組み合わせ情報 1 4 2 は、組み合わせられたセグメントに関連付けられたマイクロタスクの実行から受領される結果に基づいて、そのタスクについての最終作業生成物を構築するために、TPMS 1 2 8 によって使用される。

## 【 0 0 8 6 】

組み合わせられたセグメントが生成され、組み合わせ情報 1 4 2 が記憶されたのち、組み合わせ器サブシステム 1 2 0 は組み合わせられたセグメント 1 4 0 をさらなる処理のためにマイクロタスク生成器サブシステム 1 2 2 に転送してもよい。たとえば、組み合わせ器サブシステム 1 2 0 は、図 8 A に描かれる組み合わせられたセグメント 8 2 0、8 2 2、8 2

10

20

30

40

50

4 および 8 2 6 を、さらなる処理のためにマイクロタスク生成器サブシステム 1 2 2 に与えてもよい。

【 0 0 8 7 】

マイクロタスク生成器サブシステム 1 2 2 は、各組み合わせられたセグメントについての一つまたは複数のマイクロタスクを決定するよう構成される。ある実施形態では、マイクロタスク生成器サブシステム 1 2 2 は、組み合わせられたセグメントについて前記一つまたは複数のマイクロタスクを決定するために、マイクロタスク規則 1 4 4 を使ってもよい。マイクロタスク規則は、条件と、該条件が満たされたときに（あるいは該条件が満たされないときに）組み合わせられたセグメントに関連付けられるべき一つまたは複数のマイクロタスクとを同定してもよい。マイクロタスク規則についての条件は、実行されるべきタスク、組み合わせられたセグメントの内容、タスク要求者の素性およびその他の基準といった一つまたは複数の基準に基づいていてもよい。

10

【 0 0 8 8 】

図 8 A に描かれた例を参照するに、各組み合わせられたセグメントについて一つずつ、四つのマイクロタスクが決定されている：（ 1 ）マイクロタスク「MT1：コンピュータ描画への変換」は組み合わせられたセグメント 8 2 0 に関連付けられている；（ 2 ）マイクロタスク「MT2：タイプ・テキストへの変換」は組み合わせられたセグメント 8 2 2 に関連付けられている；（ 3 ）マイクロタスク「MT3：タイプ・テキストへの変換」は組み合わせられたセグメント 8 2 4 に関連付けられている；（ 4 ）マイクロタスク「MT4：タイプ・テキストへの変換」は組み合わせられたセグメント 8 2 6 に関連付けられている。図 8 A に描かれている例では各組み合わせられたセグメントに関連付けられているのは一つのマイクロタスクだけであるが、代替的な実施形態では、複数のマイクロタスクが単一の組み合わせられたセグメントに関連付けられてもよい。

20

【 0 0 8 9 】

図 1 に戻って参照するに、ひとたび一組の組み合わせられたセグメントについてのマイクロタスクが決定されたら、マイクロタスク生成器サブシステム 1 2 2 は、生成されたマイクロタスクについての価格 / 費用を決定するために値付けサブシステム 1 2 4 のサービスを使ってもよい。HIT は典型的には、コンピュータまたは機械によって実行されるタスクよりも、単位時間当たりより高い価格点で値付けされる。値付けサブシステム 1 2 4 は、マイクロタスクについての入力における内容の量といったさまざまな基準に基づいて、マイクロタスクについての価格点を計算するよう構成されていてもよい。たとえば、マイクロタスクは、該マイクロタスクに対応する組み合わせられたセグメント中の内容の量に基づいて値付けされてもよい。他の実施形態では、コンピュータによって実行できるマイクロタスクは人間によって実行できるマイクロタスクから区別されてもよい。たとえば、マイクロタスクの値付けに先立って、機械によって実行されるべきマイクロタスクと、人間によって実行されるべきマイクロタスクとの区別をするために前処理器 1 2 6 が使われる。ある実施形態では、前処理器サブシステム 1 2 6 は、組み合わせられたセグメントの内容を前処理して、その組み合わせられたセグメントに関連付けられたマイクロタスク（単数または複数）を実行するためにより助けとなる形にしてもよい。マイクロタスクの値付けおよび前処理器サブシステム 1 2 6 の使用に関するさらなる詳細は後述する。図 8 A に描かれた例を参照するに、価格 P1 がマイクロタスク MT1 について決定され、価格 P2 がマイクロタスク MT2 について決定され、価格 P3 がマイクロタスク MT3 について決定され、価格 P4 がマイクロタスク MT4 について決定される。

30

40

【 0 0 9 0 】

マイクロタスク生成器サブシステム 1 2 2 は、マイクロタスクおよび関連付けられた情報を、一つまたは複数のプロバイダー 1 1 4 に分配するために分配システム 1 0 6 に転送するよう構成される。マイクロタスク生成器サブシステム 1 2 2 はまた、分配システム 1 0 6 に転送されたマイクロタスクに関するマイクロタスク情報 1 4 6 を記憶してもよい。マイクロタスクについて、マイクロタスク情報 1 4 6 は、そのマイクロタスクを同定する情報、そのマイクロタスクのために関連付けられた値付け情報、そのマイクロタスクをそ

50



の入力された組み合わせられたセグメントにマッピングする情報、そのマイクロタスクが転送される先の分配システム（本質的にはMMS 104が複数の分配システムを使用してもよい実施形態において）およびその他の情報を含んでいてもよい。

【0091】

上述したように、マイクロタスク生成器サブシステム122は、マクロタスクおよび関連付けられた情報を、一つまたは複数のプロバイダー114に分配するために、分配システム106に転送するよう構成される。マイクロタスクに関連付けられた情報は、そのマイクロタスクを実行するための入力として使われる内容をもつ組み合わせられたセグメント、そのマイクロタスクについての値付け情報およびその他の情報を含んでいてもよい。

【0092】

マイクロタスクに関連付けられた情報はまた、そのマイクロタスクの実行のためのコンテキストを与えうるコンテキスト情報をも含んでいてもよい。このコンテキスト情報は、マイクロタスクの実行に関して助けとなるべくプロバイダーに提供されてもよい。たとえば、単語画像をテキスト・タイプに変換することに関わるマイクロタスクについて、マイクロタスクの精度を高めるために、そのマイクロタスクについて、入力単語画像が医療書式または名刺から抽出されたことを示すコンテキスト情報が提供されてもよい。もう一つの例として、マイクロタスクについてのコンテキスト情報は、作業者が数字、電子メール・アドレスなどをタイプする必要があるといった、そのマイクロタスクに関係するさらなる情報を提供してもよい。このように、マイクロタスクとともにプロバイダーに転送されるコンテキスト情報は、そのマイクロタスクを実行するためのコンテキストを与える情報

10

20

【0093】

ある実施形態では、MMS 104は、一組のマイクロタスクに関連付けられるべき一つまたは複数の制約条件150を決定するよう構成されていてもよい。制約条件150は、個々のマイクロタスクに関係した制約条件および/またはマイクロタスクがどのように分配されるべきかに関する制約条件を含んでいてもよい。制約条件は、マイクロタスクがどのように実行されるべきか、マイクロタスクを実行するために許容される作業者の特性、マイクロタスクについて完了までの時間の期待、どこでマイクロタスクが実行できるか（たとえば位置制約条件）、マイクロタスクについての所望される精度、分配制約条件などに関する制約条件を含んでいてもよい。作業者の特性に関係した制約条件は、たとえば、

30

【0094】

タスクに関連付けられた受け容れ可能なリスクおよび/または品質レベルは、マイクロタスク生成器サブシステム122によって生成されるマイクロタスクに関連付けられる制約条件を制御してもよい。たとえば、受け容れ可能な限度内でリスクを下げるために、特定のマイクロタスク（または一組のマイクロタスク）を、人間の作業者の代わりに、一または複数の機械作業者にアウトソーシングするほうがよいことがありうる。他方、望ましい品質の出力を得るために、特定のマイクロタスク（または一組のマイクロタスク）を一または複数の人間の作業者にアウトソーシングするほうがよいことがありうる。したがって、制約条件150は、特定のマイクロタスクが人間のプロバイダーだけに分配されるべきか、機械のプロバイダーだけに分配されるべきか、そのタスクに関連付けられたリスクおよび品質レベルに基づいて人間または機械のプロバイダーのいずれにもソーシングできるかを指定してもよい。リスクおよび/または品質因子とマイクロタスク制約条件との間のこれらの相関は、マイクロタスク規則情報144においてエンコードされていてもよく、要求されるタスクに対応する一組のマイクロタスクおよびもしあれば該一組のマイクロタスクに関連付けられるべき制約条件を決定する際に、マイクロタスク生成器サブシステム122によって使用されてもよい。

40

【0095】

50

制約条件 150 は、分配システム 106 が一組のマイクロタスクを個々のプロバイダーに分配またはアウトソーシングする仕方に関係した分配制約条件をも含んでいてもよい。たとえば、一組のマイクロタスクについての分配制約条件は、プロバイダーは該一組のマイクロタスクからの二つ以上のマイクロタスクを割り当てられないことができないことを指定してもよい。そのような制約条件は、本質的に、プロバイダーが一組のマイクロタスクから一つのマイクロタスクのみを割り当てられることができることを保証する。このことは、プロバイダーがその一つのマイクロタスクに対応する高々一つの組み合わせられたセグメントを開示されることを保証し、それによりそのプロバイダーに暴露されるのがそのタスクについて受領された入力情報の部分集合のみであることを保証するので、このことはプライバシーの保護のために重要である。

10

## 【0096】

分配制約条件 150 はまた、マイクロタスク・アウトソーシングに対する地理的制約条件を課す制約条件のような他の制約条件をも含んでいてもよい。たとえば、一組のマイクロタスクについての分配制約条件は、その一組のマイクロタスクからの二つのマイクロタスクが同じ都市内のプロバイダーに割り当てられるべきではないことを指定してもよい。これは、プロバイダー間の地理的な距離を加え、それにより入力情報のプライバシーが危殆化される可能性をさらに減らす。そのような分配は、入力情報の内容のプライバシーが危殆化されるリスクをさらに減らす（またはほとんどなくす）。

## 【0097】

いくつかの事例では、マイクロタスクについての入力情報の一つまたは複数の部分は、情報のプライバシーを保持するために編集（たとえば黒塗り〔ブラックアウト〕）されてもよい。編集されるべき領域は、タスク要求者から受領される情報に基づいて MMS 104 の人間のオペレーターによって手動でマークされてもよい。あるいはまた、編集されるべきセクションは、たとえば光学文字認識（OCR）技法、キーワード検索（タスク要求者によってプライベートであると同定された社会保障番号の検索など）などを使うことによって、自動的に決定されてもよい。

20

## 【0098】

分配システム 106 は、MMS 104 から一組のマイクロタスク（およびもしあれば関連付けられた制約条件）を受領し、それらのマイクロタスクを実行するための一または複数の作業員またはプロバイダーを決定し、それらのマイクロタスクを決定されたプロバイダーに分配するよう構成される。アウトソーシングという用語は一般に、一または複数のプロバイダーへのタスクの分配を指すのに使われる。プロバイダーは、人間の作業員および/または自動化されたコンピュータ・システム（たとえば図 1 に描かれたシステム 110）を含んでいてもよい。いくつかの実施形態では、分配システム 106 はマイクロタスクおよび関連付けられた情報を、作業員またはプロバイダーによって使用されるシステム 108（たとえばコンピュータ）に転送してもよい。プロバイダーに提供されるマイクロタスクに関連付けられた情報は、該マイクロタスクを実行するための入力として使われる一つまたは複数の組み合わせられたセグメントの内容、可能性としてはそのマイクロタスクについて決定された値付け情報、そのマイクロタスクについての期待される品質情報、そのマイクロタスクを完了するための時間枠およびその他の情報を含んでいてもよい。

30

40

## 【0099】

マイクロタスクおよびその関連付けられた情報を送達するために種々の技法が使われてもよい。場合によっては、マイクロタスクおよび関連付けられた情報は、そのマイクロタスクを実行するために選択された人間の作業員のシステムに、あるいはそのマイクロタスクを実行するシステム/機械に提供されてもよい。たとえば、実行されるべきマイクロタスクを同定する電子メールがプロバイダーに送られてもよく、そのマイクロタスクについての入力情報（すなわち、そのマイクロタスクについての組み合わせられたセグメント）がその電子メールに添付されてもよい。他の実施形態では、該情報は人間の作業員に直接与えられてもよい。ある実施形態では、分配システム 106 は、分配プロセスを容易にするために分配規則 148 を使ってもよい。ある実施形態では、アマゾン・メカニカル・ター

50

クのような分配システムは、分配システム 106 によって提供される機能を提供するよう向上されてもよい。

【0100】

分配システム 106 は、マイクロタスクが、該マイクロタスクに関連付けられた制約条件があればそれに従って分配されることを保証するよう構成される。特に、所望されるリスクまたは品質を、マイクロタスクの分配に影響するために使うことができる。そのような制約条件の結果として、一組のマイクロタスクからのマイクロタスクは、異なる地理的位置、異なるIDをもつ作業員、異なる年齢グループの作業員、異なるタイム・ゾーンの作業員、異なるアウトソーシング会社に勤める作業員などにアウトソーシングまたは分配されてもよい。

10

【0101】

分配システム 106 は、実行されるべきマイクロタスクについて一つまたは複数のプロバイダー 114 を選択するよう種々の技法を使ってもよい。ある実施形態では、最低入札値をもつプロバイダーにマイクロタスクが分配される入札システムを使ってもよい。ある実施形態では、入札システムにおいて情報のプライバシーを保護するために追加的施策が講じられてもよい。たとえば、特定のマイクロタスクについて、プロバイダーからの入札を得るために、分配システム 106 は自動的に「代表的」マイクロタスク（すなわち、たとえばタイプ・テキストへの変換のための同じ長さの語、同じ分類子信頼性（classifier confidence）など、目標の特定のマイクロタスクと同じ難しさが架空の入力内容をもつマイクロタスク）を生成してもよい。分配システム 106 は、次いで、入札に基づいて特定のプロバイダーを選択し、次いで実際の特定のマイクロタスクおよびその関連付けられた入力された組み合わせられたセグメント（単数または複数）を「落札した」入札者に分配してもよい。そのような技法を使えば、一の選択されたプロバイダーだけがマイクロタスクおよびその関連付けられた入力内容にアクセスをもつ。これは、分配プロセスのセキュリティを高める。代表的マイクロタスクを提供し、該代表的問題に対する入札に基づいてプロバイダーを選択するプロセスは、自動化されていてもよいし、あるいは何らかの人間の入力を含んでいてもよい。

20

【0102】

もう一つの実施形態では、潜在的プロバイダーが、代表的マイクロタスクを解くよう求められてもよく、その問題を解くことができる、あるいはその問題を所望される品質で所望された時間枠内に解くことができる一のプロバイダーだけが、実行されるべき実際のマイクロタスク（単数または複数）へのアクセスを得ることが許容され、他のプロバイダーは閉め出される。そのようなアプローチも、プロバイダーの素性がわかるので、セキュリティを高める、あるいはリスクを減らし、トラッキング〔追跡性〕を高める。そのようなシナリオでは、マイクロタスクに関連付けられた内容が公に暴露される（または危殆化される）場合、その情報を漏洩したプロバイダーの素性が簡単に決定できる。

30

【0103】

マイクロタスクが分配される作業員がそのマイクロタスクを受け容れないことが可能であってもよい。たとえば、作業員は、マイクロタスクに関連付けられたコスト/価格制約条件を受け容れないことがありうる。別のシナリオでは、マイクロタスクについて作業員を見出すことさえできないこともありうる。たとえば、マイクロタスクに関連付けられた作業員に関係した、地理に関係したなどの制約条件がある場合、そうした制約条件を満たす作業員を見出すことは常に可能ではないことがありうる。結果として、マイクロタスクは分配されないままとなりうる。そのようなシナリオをカバーするため、各マイクロタスクにタイムアウト値〔有効期限値〕が関連付けられていてもよい。マイクロタスクが、該マイクロタスクに関連付けられたタイムアウト値以内にアウトソーシングされることができない場合（これは作業員によってマイクロタスクが拒否された、適切な作業員が見つからなかった、その他の理由のためでありうる）、そのマイクロタスクに関係したさまざまな動作が、タイムアウトの期限切れに際してトリガーされうる。ある実施形態では、タイムアウトに際して、そのタスクが作業員によって拒否されたことが判別される場合、その

40

50

マイクロタスクは異なる作業者にアウトソーシングされてもよいし、あるいは、そのマイクロタスクは同じ作業者に、その作業者がそのマイクロタスクを受け容れる可能性を高めるような修正された制約条件をもって（たとえばより高いコスト/価格制約条件をもって）再び分配されてもよい。そのマイクロタスクについて作業者がみつからなかったために起こったタイムアウトのシナリオでは、そのタスクがより大きな、より手にはいりやすい作業者の集合に分配できるよう、そのマイクロタスクに関連付けられた制約条件が変更（典型的には低下）されてもよい。このようにして、マイクロタスクに関連付けられたタイムアウトに際して、そのマイクロタスクを再分配するためにさまざまな動作が実行されてもよい。

**【 0 1 0 4 】**

分配システム 1 0 6 はまた、プロバイダー 1 1 4 からマイクロタスク実行の結果または作業生成物（マイクロタスク生成物と称する）を受け取るよう構成される。マイクロタスク生成物は、一または複数の作業者システム 1 0 8 からおよび/または一または複数の自動化されたシステム 1 1 0 から受領されてもよい。分配システム 1 0 6 は、マイクロタスク生成物を MMS 1 0 4 に転送するよう構成される。いくつかの実施形態では、一または複数のプロバイダーは、マイクロタスク生成物を MMS 1 0 4 に直接提供してもよい。

**【 0 1 0 5 】**

ある実施形態では、分配システム 1 0 6 は、マイクロタスク生成物を受領するために、プロバイダー 1 1 4 のシステムをポーリングしてもよい。ある代替的な実施形態では、プロバイダー・システムが、マイクロタスク生成物を分配システム 1 0 6 にプッシュするよう構成されてもよい。さらに、ある実施形態では、MMS 1 0 4 は、マイクロタスク生成物を受領するために分配システム 1 0 6 にポーリングしてもよい。一方、他の実施形態では、分配システム 1 0 6 はマイクロタスク生成物を MMS 1 0 4 にプッシュするよう構成されていてもよい。

**【 0 1 0 6 】**

一または複数の分配システム 1 0 6 から MMS 1 0 4 によって受領されるマイクロタスク生成物は、タスク生成物管理サブシステム 1 2 8（TPMS）に転送される。TPMS 1 2 8 は、そのタスクに対応する諸マイクロタスクについて受領されたマイクロタスク生成物に基づいて、そのタスクについての最終作業生成物を構築するよう構成される。ある実施形態では、そのタスクについての最終生成物は、マイクロタスクについて受領されたマイクロタスク作業生成物を総合することによって生成されてもよい。TPMS 1 2 8 は、総合を実行するためのアセンブラー・モジュール 1 2 9 を有していてもよい。次いで、最終生成物は、タスク要求者に提供されてもよい。

**【 0 1 0 7 】**

ある実施形態では、TPMS 1 2 8 は、タスクに対応するマイクロタスクについて受領されたマイクロタスク生成物に基づいて、そのタスクについての最終作業生成物を構築するために、マイクロタスク情報 1 4 6、組み合わせ情報 1 4 2 およびセグメント分割情報 1 3 4 を使う。たとえば、分配システム 1 0 6 から受領されたマイクロタスク生成物について、TPMS 1 2 8 は、マイクロタスク生成物に対応するマイクロタスクおよびそのマイクロタスクに対応する組み合わせられたセグメント（すなわち、そのマイクロタスクについて入力として使われた組み合わせられたセグメント）を決定するためにマイクロタスク情報 1 4 6 を使ってもよい。このようにして、TPMS は、マイクロタスク情報 1 4 6 を使って、受領された各マイクロタスク生成物を組み合わせられたセグメント（単数または複数）にマッピングしてもよい。TPMS 1 2 8 は、組み合わせられたセグメントにマッピングされた一または複数のマイクロタスク生成物に基づいて、各組み合わせられたセグメントについて作業生成物を生成してもよい。TPMS 1 2 8 は次いで、組み合わせ情報 1 4 2 を使って、組み合わせられたセグメントに対応するセグメントを決定してもよい。TPMS 1 2 8 は次いで、各セグメントについて、該セグメントに対応する組み合わせられたセグメントについての作業生成物に基づいて、作業生成物を構築してもよい。TPMS 1 0 4 は次いで、それらのセグメントを、そのタスクについて受領された入力情報における個々の入力文書にマッピン

10

20

30

40

50

グするために、セグメント分割情報 1 3 4 を使ってもよい。TPMS 1 2 8 は、セグメント分割情報 1 3 4 を、セグメントについて構築された作業生成物に基づいて各入力文書についての作業生成物を構築してもよい。入力文書について構築された作業生成物は、そのタスクについての最終作業生成物を表していてもよい。

【 0 1 0 8 】

先に論じたように、一つのマイクロタスク（またはその一部）についての作業生成物または出力は、別のマイクロタスクについての入力情報として使われてもよい。たとえば、機械が第一のマイクロタスクを実行することによって生成される作業生成物は、人間の作業員によって実行されるべき第二のマイクロタスクについての入力として使われてもよい。したがって、第一のマイクロタスク・タスクをある「作業員」に提出し、該マイクロタスクを実行することから得られた結果を受領し、第一のマイクロタスク実行から受領された結果（またはその一部）を入力とする新たなマイクロタスクを生成することが可能である。ある実施形態では、第一のマイクロタスクから受領された結果は、他の情報（たとえば、別のマイクロタスクから受領された結果）と組み合わせられてもよく、組み合わせられた情報が、別の作業員または自動化システムにソーシングされる新しいマイクロタスクについての入力として使われてもよい。別の実施形態では、第一のマイクロタスクから受領された結果は部分集合にセグメント分割され、各部分集合について新しいマイクロタスクが決定され、次いで新しい一組のマイクロタスクが実行されるために作業員に送られてもよい。

【 0 1 0 9 】

したがって、ある種の事例では、TPMS 1 2 8 は受領されたマイクロタスク生成物の一または複数をマイクロタスク生成器サブシステム 1 2 2 に転送してもよい。TPMS 1 2 8 からマイクロタスク生成物を受領するに際して、マイクロタスク生成器サブシステム 1 2 2 は、一または複数のマイクロタスクからなる新たな組を生成してもよい。ここで、受領されたマイクロタスク生成物が新しいマイクロタスクについての入力である。これらの新しいマイクロタスクは次いで、値付けサブシステム 1 2 4 を使って値付けされてもよく、次いで一または複数のプロバイダーへの分配のために分配システム 1 0 6 に送られてもよい。

【 0 1 1 0 】

ある実施形態では、マイクロタスクについてプロバイダーから受領されたマイクロタスク生成物の品質が検査されてもよく、必要な品質閾値を満たさないと判定された場合、マイクロタスクは別のプロバイダーへの分配のために分配システム 1 0 6 に再提出されてもよい。たとえば、TPMS 1 2 8 は、オーディオ・セグメントの文字起こしから帰結するマイクロタスク生成物を受領してもよい。TPMS 1 2 8 は次いで、そのマイクロタスク生成物に関連付けられた信頼性スコアを決定してもよい。信頼性スコアが何らかのユーザー構成設定可能な閾値を下回る場合、TPMS 1 2 8 は、文字起こしをやり直す必要があると判定してもよく、そのマイクロタスク生成物をマイクロタスク生成器サブシステム 1 2 2 に送ってもよい。マイクロタスク生成器サブシステム 1 2 2 は次いで、そのマイクロタスク生成物に対応する（オーディオ情報を含む）組み合わせられたセグメントを決定し、その組み合わせられたセグメントについての新たな文字起こしマイクロタスクを生成してもよい。新しいマイクロタスクは、最初にそのマイクロタスクを実行したプロバイダー以外のプロバイダーへの分配のために、分配システム 1 0 6 に送られてもよい。

【 0 1 1 1 】

図 8 B は、図 8 A において生成され複数のプロバイダーに分配されたマイクロタスクに対応する受領されたマイクロタスク生成物に基づいて、図 8 A において受領されたタスクについて、いかにして最終作業生成物が構築されるかの例を示している。図 8 B に描かれているように、マイクロタスク MT1、MT2、MT3 および MT4 に対応するマイクロタスク生成物（MTP: microtask product）が MMS 1 0 4 によって受領されてもよい。マイクロタスク MT1 に対応するマイクロタスク生成物 8 4 0 は作業員 1 から受領され、組み合わせられたセグメント 8 2 0 に含まれるグラフィックに対応するコンピュータ描画を含む。マイクロタ

スクMT2に対応するマイクロタスク生成物 8 4 2 は作業員 2 から受領され、組み合わされたセグメント 8 2 2 に含まれるさまざまなセグメントの内容に対応するタイプ・テキストを含む。マクロタスクMT3に対応するマイクロタスク生成物 8 4 4 は作業員 3 から受領され、組み合わされたセグメント 8 2 4 に含まれるさまざまなセグメントの内容に対応するタイプ・テキストを含む。マクロタスクMT4に対応するマイクロタスク生成物 8 4 6 は作業員 4 から受領され、組み合わされたセグメント 8 2 6 に含まれるさまざまなセグメントの内容に対応するタイプ・テキストを含む。

【 0 1 1 2 】

ある実施形態では、受領された各マイクロタスク生成物は、使用された作業員および訂正のシーケンスまたはマイクロタスクに対して実行された検査に依存する、関連付けられた品質推定値 $Q_1$ 、 $Q_2$ 、 $Q_3$ 、 $Q_4$ を持っていてもよい。この品質は、認識されたカードの品質ではなく、マイクロタスクについての所望される結果への推定される近さである。

10

【 0 1 1 3 】

マイクロタスク情報 1 4 6 および組み合わせ情報 1 4 2 を使って、TPMS 1 2 8 は次いで、マイクロタスク生成物をその対応する組み合わせられたセグメントに、最終的にはセグメント 8 0 8 および 8 1 4 にマッピングしてもよい。図 8 B に描かれているように、マイクロタスク生成物 8 4 0 の内容の一部 (XYZロゴのコンピュータ描画) がセグメント 8 1 0 にマッピングされる。これがセグメント 8 1 0 についての生成物を表す。マイクロタスク生成物 8 4 0 の内容の一部 (ABCロゴのコンピュータ描画) がセグメント 8 1 6 にマッピングされる。これがセグメント 8 1 6 についての生成物を表す。同様に、8 1 2 および 8 1 8 におけるセグメントについての生成物が、マイクロタスク生成物 8 4 2、8 4 4 および 8 4 6 の内容から決定される。8 1 2、8 1 6 および 8 1 8 における各セグメントについての生成物は、セグメント内容に対応するタイプ・テキスト部分のリストを含む。

20

【 0 1 1 4 】

TPMS 1 2 8 は次いで、それらのセグメントを、そのタスクについて受領された入力情報における個々の入力文書 8 0 4 および 8 0 6 にマッピングし、それらの入力文書についての作業生成物を構築する。TPMS 1 2 8 は、セグメントをタスクについての個々の入力文書にマッピングするためにセグメント分割情報 1 3 4 を使ってもよい。図 8 B に描かれるように、文書 8 0 4 についての最終作業生成物は、文書 8 0 4 内のテキスト部分に対応するタイプ・テキストおよび文書 8 0 4 内のXYZロゴ・グラフィックに対応するコンピュータ描画を含む。文書 8 0 6 についての最終作業生成物は、文書 8 0 6 内のテキスト部分に対応するタイプ・テキストおよび文書 8 0 6 内のABCロゴ・グラフィックに対応するコンピュータ描画を含む。このように、タスク 8 0 0 について最終作業生成物は、そのタスクについての入力情報の一部として受領された各入力文書について構築される作業生成物を含む。

30

【 0 1 1 5 】

ある実施形態では、最終作業生成物について品質推定値が与えられる。この品質推定値は、個々のマイクロタスク生成物に関連付けられた品質推定値に依存する。最も単純な場合では、最終作業生成物についての品質推定値は、その最終作業生成物を形成するために総合されたマイクロタスク作業生成物に関連付けられた品質推定値の平均であってもよい。あるいはまた、該品質推定値は、最終作業生成物における、各マイクロタスクの一部であった項目の数に依存して重み付けされてもよい。

40

【 0 1 1 6 】

図 1 に戻って参照すると、タスクについての最終作業生成物を構築すると、TPMS 1 2 8 は、タスク要求者に最終作業生成物を転送するよう構築されてもよい。最終作業生成物は、タスク要求者に、いくつかの異なる方法を使って転送されうる。ある実施形態では、作業生成物は、タスク要求者によって指定された、最終作業生成物を記憶するためのメモリ位置に記憶されてもよい。最終作業生成物はまた、タスク要求者に通信されてもよい。たとえば、図 8 B に描かれる最終作業生成物はWORD〔ワード〕文書に書き込まれてもよく

50

、そのWORD文書が次いでタスク要求者に通信されてもよい。最終作業生成物に対して、他のユーザー構成設定可能な動作がMMS 104によって実行されてもよい。

【0117】

ある実施形態では、マイクロタスク生成器サブシステム122によって決定された一組のマイクロタスクは、複製されたマイクロタスクを含んでいてもよい。たとえば、文字起こしされるべきオーディオ情報を含む組み合わせられたセグメントについて、文字起こしの精度を高めるために、同じ組み合わせられたセグメントに関連付けて、それぞれがそのオーディオ情報が文字起こしされるべきことを指定する複数の重複したマイクロタスクが生成されてもよい。それら複数の重複したマイクロタスクには、互いに異なるプロバイダーに分配されるという制約条件が関連付けられてもよい。すると、MMS 104は、該異なるプロバイダーから重複したマイクロタスクに対応して受領されるマイクロタスク生成物どうしを比較して、文字起こしの精度を決定してもよい。

10

【0118】

別のシナリオでは、マイクロタスクおよび関連付けられた入力情報は、同じプロバイダーに複数回アウトソーシングされ、それによりそのマイクロタスクが複数回実行されてもよい。そのマイクロタスクが複数回実行されることから帰結する複数のマイクロタスク生成物は、結果として得られる作業生成物の品質を向上させることがありうる。

【0119】

たいていの場合は（たとえば90%の場合）コンピュータがうまく行すが、残りの場合についてはひどく失敗するタスクがたくさんある。この品質レベルはしばしば十分ではなく、通常、これらのタスクは人間に与えられる。そのようなシナリオにおいて、MMS 104は、コンピュータ・システムによる実行のために第一の組のマイクロタスクを生成してもよい。第一の組のマイクロタスクについて得られた結果に基づいて、MMS 104は、人間の作業者に分配される第二の組のマイクロタスクを生成してもよい。ここで、第一の組のマイクロタスクの実行から得られた結果が、第二の組のマイクロタスクへの入力として使われる。第二の組のマイクロタスクは、第一の組のマイクロタスクから得られる結果における誤りを訂正することを含んでいてもよい。このようにして、人間が、それ以外の点では自動化されたプロセスによってなされた誤りを訂正するために使われてもよい。全体的なタスクの品質およびその品質の推定値は、人間によってなされるおよび機械によってなされる作業の品質の複雑な組み合わせである。

20

30

【0120】

もう一つの実施形態では、あるタスクについて二組のマイクロタスクが生成されてもよい。第一の組は機械による実行にとってより好適であり、第二の組は人間による実行にとってより好適である。人間により実行されるマイクロタスクとコンピュータにより実行されるマイクロタスクの組み合わせがこのようにして、半自動化された仕方でタスクを効率的に解くために使用されてもよい。このハイブリッド・モデルはいくつかの恩恵をもたらす。たとえば、人間は、プロセスの自動化された部分について、コンピュータによる誤りを少なくすることができるトレーニング・データを提供することができる。さらに、より平凡なマイクロタスクは機械/コンピュータにアウトソーシングされる一方、より高い認知要件をもつマイクロタスクは人間にアウトソーシングされてもよい。これにより、人間の作業者にとっての作業がより興味深いものとなり（そのことはよりよい品質につながりうる）、一方、タスク全体を完成させるために必要とする全体的な人間時間はより少なくなる。機械によって実行されるマイクロタスクに関連付けられるコストは人間によって実行されるマイクロタスクよりも概して安価なので、これは、全体的なタスクの費用を下げることもありうる。

40

【0121】

品質管理は、どんなアウトソーシング・モデルでもきわめて重要である。プロバイダー、特に人間の作業者は、評価され、作業の品質についてのフィードバックを与えられる必要がある。伝統的に、これは、タスク（たとえばマイクロタスク）の出力を吟味し、フィードバックを提供する人間によってなされる。しかしながら、このプロセスは、特にタス

50

クが複数のより小さなマイクロタスクに分解され、そのそれぞれが別個に評価される必要があるときには、タスク自身と同じくらい高価になることがある。ある実施形態では、MMS 104は、タスクについての品質管理を自動化するよう構成される。それは、品質管理を、それ自身マイクロ・アウトソーシングされることのできる（よって人間によって実行されるものおよびコンピュータによって実行されるものを含むより小さなジョブに分解される）マイクロタスクとして成形することによる。マイクロタスクの該生成において品質を上げるいくつかの方法がある。一つの方法は、複数の作業者に同じタスクを実行させて、二の作業者が一致するときのみ結果を受け容れることである。これは品質を高める。 $Q(I, W_1, W_2) > Q(I, W)$ 。そのような技法は、複数の作業者がタスク全体を行うことを要求するので、より高価となりうる。

10

## 【0122】

もう一つの技法によれば、入力の中から部分集合が、何らかの重なりをもって、複数の作業者に送られることができる。たとえば、 $I = I_1 + I_2 + I_3$ であれば、 $W_1$ は $I_1$ および $I_2$ に対して作業でき、 $W_2$ は $I_2$ および $I_3$ に対して作業できる。よって、どの作業者も入力全体を見ることはなく、リスクを低下させ、それでいて各入力は一以上の作業者によって処理される。ある実施形態では、品質に関係した動作に関するマイクロタスクを生成することによって、品質管理がタスクと統合されることができる。たとえば、書式における数値項目を認識することからなるタスクであって、書式には合計値も記入されており、これも認識すべきであり、転写〔認識〕された個々の値の和が前記合計値の転写になるべきである場合を考える。転写された値は自動化タスクによって簡単に合計でき、計算された和が転写された和と一致すれば、転写の品質は良好であると考えられることができる。それらの値が一致しなければ、品質を検査するために追加的なマイクロタスクが必要となることありうる。それはたとえば、作業者に、もとの書式の合計が正しく計算されていたことを検証させる、あるいはタスクの一部を異なる作業者をもって繰り返すことによる。

20

## 【0123】

マイクロタスクの実行をより興味深いものとするとともに品質を改善するもう一つのアプローチは、作業者の実績を何らかのフォーマットで公開することを含む。たとえば、作業プロバイダーからのフィードバックを使って何らかのスコアを確立することができ、スコアが上位の作業者が、現在ゲームやソーシャル・ネットワークで行われているように、「高スコア」リスト上で公開されることができる。作業プロバイダー・フィードバックに加えて、作業者は、タスク完了のスピードまたは実行されたタスクの多様さについて認められることができる。何らかの品質尺度において上位の作業者にはボーナスが与えられてもよい。

30

## 【0124】

品質管理に対するもう一つのアプローチは、自己報告品質に関わる。しばしば、人間の作業者は、該作業者によって実行されたマイクロタスクの結果に対する自信/信頼性を正確に報告できる。たとえば、文字起こし/転写のマイクロタスクについて、そのマイクロタスクを実行する人間の作業者は、そのマイクロタスクの実行に対するフィードバック（たとえば信頼性スコア）を与える能力を備えていることがある。このフィードバックは、そのマイクロタスクをやり直す必要があるかどうかを決定する際に、MMS 104によって使用されてもよい。作業者（worker）による品質の推定値 $Q_{\text{worker}}(I)$ は、所望される品質レベルと比較されることができ、所望される品質が達成されていれば、タスクは完了と考えられる。所望される品質が達成されていなければ、品質を改善するために訂正マイクロタスクが使われてもよいし、あるいはその作業が別のマイクロタスクを使って別の作業者によって繰り返されたり別の作業者からの出力と組み合わせられたりしてもよい。作業者は、ある種のジョブをもっと得ること、あるいは何らかのジョブもしくは何らかの型のジョブについて追加的なトレーニングを得ることに関心があることを報告することができてよい。これは、作業プロバイダーが追加的な指示を与え、品質を改善することを許容しうる。作業者からの、自らの作業に対する信頼性の報告は、自動的な統計とともに、特定のジョブについてのジョブ品質を推定し、新しいタスクの割り当てを決定するために使う

40

50



ことができる。

【0125】

MMS 104は図1に描かれている実施形態では単一のシステムとして示されているが、代替的な実施形態では、MMS 104によって実行される機能は、互いと協働する複数のシステムによって実行されてもよい。さらに、MMS 104および分配システム106が図1では別個のシステムとして示されているが、代替的な実施形態では、MMS 104および分配システム106によって実行される機能は単一のシステムまたは複数のシステムによって実行されてもよい。したがって、図1に描かれている実施形態は、請求項に記載される本発明の範囲を限定することを意図したものではない。他の変形が可能である。

【0126】

図2は、本発明のある実施形態に基づく、タスクについて受領される入力情報の内容のプライバシーを保持しつつ、該タスクを実行する高レベルの方法を記述する簡略化されたフローチャート200を描いている。図2に描かれる処理は、プロセッサによって実行されるソフトウェア(たとえばプログラム、コード、命令)、ハードウェアまたはそれらの組み合わせによって実行されてもよい。ソフトウェアは、コンピュータ可読記憶媒体上に記憶されていてもよい。図2に描かれる処理ステップの具体的な系列は、請求項に記載される本発明の実施形態の範囲を限定することを意図したものではない。

【0127】

図2に描かれるように、処理は、タスク要求を受領するに際して開始されてもよい(ステップ202)。タスク要求は、実行されるべきタスクを同定してもよく、そのタスクを実行するために使われるべき入力情報を指定してもよい。代替的な実施形態では、タスク要求は入力情報を含んでいてもよく、MMS 104は、該入力情報の属性に基づいて実行されるべきタスクを自動的に決定してもよい。たとえば、MMS 104は、実行されるべきタスクを決定するためにタスク規則130を使ってもよい。202において受領される情報はまた、受け容れ可能なリスク・レベル、所望される品質レベル、コスト閾値などといった、実行されるべきタスクについての一つまたは複数の因子または制約条件を含んでいてもよい。

【0128】

タスクについての因子または制約条件が決定される(ステップ203)。これらの因子のうちの一つまたは複数は、202において受領されるタスク要求において指定されてもよい。たとえば、タスク要求者は、タスクについての受け容れ可能なリスク閾値、タスクについての期待される品質閾値、タスクを実行するためにコストなどの一つまたは複数、タスク要求を介して指定してもよい。ある実施形態では、MMS 104は、タスク要求の解析に基づいて、そのタスクについての因子を、もしあれば決定するよう構成されてもよい。たとえば、実行されるべきタスクの性質に基づいて、および/またはそのタスクについて与えられる入力情報の特性に基づいて、MMS 104は、そのタスクに関連付けられるべき一つまたは複数の因子または制約条件の組を決定してもよい。さらにもう一つのシナリオでは、MMS 104について構成設定されたいくつかのデフォルト制約条件が決定されてもよく、要求されたタスクについて使用されてもよい。203におけるタスクについて決定された因子または制約条件は、タスク要求がどのように処理されるかに影響しうる。たとえば、204、206、208、210、211、212、214、216および218に関して後述されるさまざまな処理は、203において決定された制約条件が満たされるまたは満足されるように実行されうる。

【0129】

入力情報は次いで、一組のセグメントにセグメント分割され、ここで、各セグメントは、202において受領された入力情報の内容の一部または部分集合を含む(ステップ204)。セグメント分割は、そのタスクについて選択された一組のセグメント分割規則に基づいて実行されてもよい。受け容れ可能なリスク・レベル、所望される品質レベル、コスト閾値などといった、実行されるべきタスクに関連付けられる制約条件は、セグメント分割に影響してもよい。図1に描かれる実施形態では、204における処理はセグメント分

10

20

30

40

50

割器サブシステム 1 1 8 によって実行されてもよい。

【 0 1 3 0 】

次いで、2 0 4 において生成されたセグメントに基づいて、一組の組み合わせられたセグメントが生成されてもよい（ステップ 2 0 6）。2 0 6 において生成される組み合わせられたセグメントは、2 0 4 において生成されたセグメントのうちの一つまたは複数またはその一部を含んでいてもよい。受け容れ可能なリスク・レベル、所望される品質レベル、コスト閾値などといった、実行されるべきタスクに関連付けられる制約条件は、セグメントが組み合わせられる仕方に影響してもよい。たとえば、図 1 について先述したように、どの一つ of の組み合わせられたセグメントに対応する情報も全体的な入力情報の内容を危殆化させないよう、複数の組み合わせられたセグメントにまたがって情報が分配されてもよい。図 1

10

【 0 1 3 1 】

ある実施形態では、一組の組み合わせられたセグメントの生成は、毎回実行されなくてもよい。そのような実施形態では、組み合わせられたセグメントが生成されるかどうかは、実行されるべきタスクに関連付けられたリスク・レベルに依存してもよい。組み合わせられたセグメントの生成はセグメント分割を超えて入力情報をさらにぼかし、よってアウトソーシングに関連するリスクを減らす助けになる。したがって、ある実施形態では、組み合わせられたセグメントは、タスクに関連付けられたリスク・レベルが何らかの閾値を下回るときにのみ、生成されてもよい。たとえば、組み合わせは、タスクに関連付けられた受け容

20

【 0 1 3 2 】

次いで、2 0 6 において生成された各組み合わせられたセグメントについて、一つまたは複数のタスク（マイクロタスク）が決定される（ステップ 2 0 8）。図 1 に描かれた実施形態では、2 0 8 の処理は、マイクロタスク生成器サブシステム 1 2 2 によって実行されてもよい。

【 0 1 3 3 】

2 0 8 で決定されたマイクロタスクのうちの一つまたは複数について、値付け情報が決定されてもよい（ステップ 2 1 0）。図 1 に描かれた実施形態では、2 1 0 の処理は値付けサブシステム 1 2 4 によって実行されてもよい。2 0 8 で決定された一組のマイクロタスクの分配に影響するさらなる制約条件がもしあれば決定される（ステップ 2 1 1）。これらの制約条件は、2 0 8 において決定された一組のマイクロタスクにおける個々のマイクロタスクに係した制約条件および/またはマイクロタスクの組に適用可能な制約条件を含んでいてもよい。

40

【 0 1 3 4 】

次いで、一組のマイクロタスクは、一または複数のプロバイダーに分配（アウトソーシング）されてもよい（ステップ 2 1 2）。マイクロタスクに関連付けられた情報が、2 1

50

2の一部として分配されてもよい。マイクロタスクに関連付けられた情報は、そのマイクロタスクについての入力として使われる情報およびそのマイクロタスクに関連付けられた値付け情報を含む組み合わせられたセグメントを含んでいてもよい。いくつかの実施形態では、マイクロタスクの実行を容易にするためのツール/リソースもマイクロタスクとともに分配されてもよい。たとえば、マイクロタスクがグラフィックをコンピュータ描画に変換することに関わる場合、コンピュータ描画アプリケーション(たとえばVISIO)がマイクロタスクとともに分配/配送されてもよい。212における分配は、マイクロタスクに関連付けられており、211において決定された制約条件があればそれに従いつつ、実行されてもよい。プロバイダーは人間の作業員でも自動化されたシステムでもよい。

【0135】

マイクロタスクの実行に際して、マイクロタスクに対応する作業生成物が受領されてもよい(ステップ214)。次いで、202において受領されたタスクについての最終作業生成物が、214において受領されたマイクロタスク生成物に基づいて構築されてもよい(ステップ216)。図1に描かれる実施形態では、216の処理はTPMS 128によって実行されてもよい。次いで、216において構築されたタスクについての作業生成物に対して、動作が任意的に実行されてもよい(ステップ218)。たとえば、最終作業生成物がメモリに記憶される、タスク要求者に通信される、などしてもよい。

【0136】

図3は、本発明のある実施形態に基づく入力情報をセグメント分割するために実行される処理を記述する簡略化されたフローチャート300を描いている。図3に描かれた処理は、プロセッサによって実行されるソフトウェア(たとえばプログラム、コード、命令)、ハードウェアまたはそれらの組み合わせによって実行されてもよい。ソフトウェアは、コンピュータ可読記憶媒体上に記憶されていてもよい。図3に描かれる処理ステップの具体的な系列は、請求項に記載される本発明の実施形態の範囲を限定することを意図したものではない。ある実施形態では、図3に描かれる処理は、図2のステップ204の一部として実行されてもよく、図1のセグメント分割器サブシステム118によって実行されてもよい。

【0137】

図3に描かれるように、処理は、実行されるべきタスクについての入力情報を受領するのに際して開始されてもよい(ステップ302)。そのタスクについて使用されるべき一組のセグメント分割規則が決定される(ステップ304)。タスク要求者の素性、タスク要求の源(たとえばタスク要求の受信元のコンピュータのIPアドレス、タスク要求の受信元の地理的領域など)、入力情報の内容およびその他の因子といった、さまざまな因子が、304でセグメント分割規則を選択するために使用されてもよい。受け入れ可能なリスク・レベル、所望される品質レベル、コスト閾値などといった、実行されるべきタスクに関連付けられる制約条件は、セグメント分割規則の選択に影響してもよい。302において受領された入力情報は次いで304において決定された一組のセグメント分割規則を使ってセグメント分割されて、一つまたは複数のセグメントの組を生成する(ステップ306)。306において実行されたセグメント分割について、セグメント分割情報が記憶される(ステップ308)。306において生成された一組のセグメントは、さらなる処理のために組み合わせ器サブシステム120に与えられる。

【0138】

図4は、本発明のある実施形態に基づく組み合わせられたセグメントを生成するために実行される処理を記述する簡略化されたフローチャート400を描いている。図4に描かれた処理はプロセッサによって実行されるソフトウェア(たとえばプログラム、コード、命令)、ハードウェアまたはそれらの組み合わせによって実行されてもよい。ソフトウェアは、コンピュータ可読記憶媒体上に記憶されていてもよい。図4に描かれる処理ステップの具体的な系列は、請求項に記載される本発明の実施形態の範囲を限定することを意図したものではない。ある実施形態では、図4に描かれる処理は、図2のステップ206の一部として実行されてもよく、図1の組み合わせ器サブシステム120によって実行されて

10

20

30

40

50

もよい。

【0139】

図4に描かれるように、処理は、実行されるべきタスクについての入力情報から生成された一組のセグメントを受領するのに際して開始されてもよい(ステップ402)。組み合わせられたセグメントを生成するために使用されるべき一組の組み合わせ規則が決定される(ステップ404)。タスク要求者の素性、タスク要求の源、セグメントの内容およびその他の因子といった、さまざまな因子が、404において組み合わせ規則を選択するために使用されてもよい。受け容れ可能なリスク・レベル、所望される品質レベル、コスト閾値などといった、実行されるべきタスクに関連付けられる制約条件も、404における組み合わせ規則の選択に影響してもよい。次いで、一つまたは複数の組み合わせられたセグメントの組が、402において受領されたセグメントに基づいて、404において決定された組み合わせ規則を使って生成される(ステップ406)。各組み合わせられたセグメントは、402において受領されたセグメントの一つまたは複数またはその一部を含んでいてもよい。組み合わせ情報が記憶されてもよい(ステップ408)。他の情報に加えて、408において記憶される組み合わせ情報は、各組み合わせられたセグメントについて：組み合わせられたセグメントを同定する情報、組み合わせられたセグメントをその構成要素である一つまたは複数のセグメントにマッピングする情報(すなわち、組み合わせられたセグメントに内容が含まれている諸セグメントの情報)、組み合わせられたセグメント内の諸セグメントの位置およびその他の情報を含んでいてもよい。次いで、406において生成された一組の組み合わせられたセグメントは、さらなる処理のためにマイクロタスク生成器サブシステム122に与えられる(ステップ410)。

10

20

【0140】

先述したように、組み合わせられたセグメントが生成されるかどうかは、実行されるべきタスクに関連付けられたリスク・レベルに依存してもよい。さらに、組み合わせられたセグメントを生成するために使われる組み合わせ技法は、さまざまな異なるリスク・レベルについて異なってもよい。組み合わせられたセグメントを生成するためのリスク・レベル閾値および該さまざまなリスク・レベルに対応する使用されるべきさまざまな組み合わせ技法に関係した情報が、組み合わせられたセグメントを生成するために組み合わせ器120によって使用される組み合わせ規則138においてエンコードされてもよい。

【0141】

図5は、本発明のある実施形態に基づく、マイクロタスク生成器サブシステムによって実行される処理を示す簡略化されたフローチャート500を描いている。図5に描かれた処理はプロセッサによって実行されるソフトウェア(たとえばプログラム、コード、命令)、ハードウェアまたはそれらの組み合わせによって実行されてもよい。ソフトウェアは、コンピュータ可読記憶媒体上に記憶されていてもよい。図5に描かれる処理ステップの具体的な系列は、請求項に記載される本発明の実施形態の範囲を限定することを意図したものではない。

30

【0142】

図5に描かれるように、処理は、タスクについての一組の組み合わせられたセグメントを受領するのに際して開始されてもよい(ステップ502)。502において受領された一組の組み合わせられたセグメントについて一組のマイクロタスク規則が決定されてもよい(ステップ504)。タスク要求者の素性、タスク要求の源、組み合わせられたセグメントの内容およびその他の因子といったさまざまな因子が、マイクロタスク規則を選択するために使用されてもよい。受け容れ可能なリスク・レベル、所望される品質レベル、コスト閾値などといった、実行されるべきタスクに関連付けられる制約条件が、504におけるマイクロタスク規則の選択に影響してもよい。次いで、一つまたは複数のマイクロタスクが、502において受領された各組み合わせられたセグメントについて、504において決定されたマイクロタスク規則の一つまたは複数を使って決定される(ステップ506)。506において決定されたマイクロタスクの一つまたは複数について、値付け情報が決定されてもよい(ステップ508)。マイクロタスクについての値付け情報を決定するための

40

50

処理は、値付けサブシステム 1 2 4 によって提供されるサービスを使ってもよい。前記一組のマイクロタスクに関連付けられるべき制約条件があれば決定される（ステップ 5 1 0）。これらの制約条件は、前記一組のマイクロタスクにおける個々のマイクロタスクに関係した制約条件および/または前記マイクロタスクのプロバイダーへの分配に係した制約条件を含んでいてもよい。受け容れ可能なリスク・レベル、所望される品質レベル、コスト閾値などといった、実行されるべきタスクに関連付けられる制約条件が、5 1 0 において前記一組のマイクロタスクについて決定される制約条件に影響してもよい。

#### 【 0 1 4 3 】

次いで、前記一組のマイクロタスクは、関連付けられた情報とともに、一または複数のプロバイダーへの分配のために分配システムに転送される（ステップ 5 1 2）。前記一組のマイクロタスクに関連付けられた情報は、各マイクロタスクについて、該マイクロタスクを実行するための入力として使われるべき内容を含む組み合わせられたセグメント（単数または複数）、そのマイクロタスクについて決定された値付け情報、そのマイクロタスクについての制約条件（もしあれば）および前記一組のマイクロタスクに関連付けられた分配制約条件を含んでいてもよい。マイクロタスク生成器サブシステムはまた、前記一組のマイクロタスクについてのマイクロタスク情報を記憶してもよい（ステップ 5 1 4）。マイクロタスク情報は、各マイクロタスクに関連付けられた値付け情報、マイクロタスクをその組み合わせられたセグメントにマッピングする情報、マイクロタスクの転送先の分配システム（特にMMS 1 0 4 が複数の分配システムを使用しうる実施形態において）およびその他の情報を含む、マイクロタスクに係した情報を含んでいてもよい。

#### 【 0 1 4 4 】

図 6 は、本発明のある実施形態に基づく、分配システムによって実行される処理を示す簡略化されたフローチャート 6 0 0 を描いている。図 6 に描かれた処理はプロセッサによって実行されるソフトウェア（たとえばプログラム、コード、命令）、ハードウェアまたはそれらの組み合わせによって実行されてもよい。ソフトウェアは、コンピュータ可読記憶媒体上に記憶されていてもよい。図 6 に描かれる処理ステップの具体的な系列は、請求項に記載される本発明の実施形態の範囲を限定することを意図したものではない。

#### 【 0 1 4 5 】

図 6 に描かれるように、処理は、分配システムが一組のマイクロタスクおよび関連付けられた情報を受領するときに開始されてもよい（ステップ 6 0 2）。関連付けられた情報は、マイクロタスクに対応する組み合わせられたセグメントおよび可能性としてはまた、マイクロタスクに関連付けられた一つまたは複数の制約条件を含んでいてもよい。前記一組のマイクロタスクにおけるマイクロタスクは次いで分配システムによって分配される（ステップ 6 0 4）。6 0 4 における分配は、マイクロタスクに関連付けられた制約条件があればそれが満足されるように実行される。分配システムはまた、プロバイダーによるマイクロタスクの実行から帰結する作業生成物（マイクロタスク生成物と称される）を受領もする（ステップ 6 0 6）。次いで、マイクロタスク生成物はさらなる処理のためにマイクロタスク管理システム（図 1 に描かれるMMS 1 0 4 のような）に転送されてもよい（ステップ 6 0 8）。

#### 【 0 1 4 6 】

図 7 は、本発明のある実施形態に基づく、あるタスクに対応する諸マイクロタスクについて受領されたマイクロタスク生成物に基づいて、該タスクについての最終作業生成物を生成するために実行される処理を示す簡略化されたフローチャート 7 0 0 を描いている。図 7 に描かれた処理はプロセッサによって実行されるソフトウェア（たとえばプログラム、コード、命令）、ハードウェアまたはそれらの組み合わせによって実行されてもよい。ソフトウェアは、コンピュータ可読記憶媒体上に記憶されていてもよい。図 7 に描かれる処理ステップの具体的な系列は、請求項に記載される本発明の実施形態の範囲を限定することを意図したものではない。ある実施形態では、図 7 に描かれる処理は、図 2 のステップ 2 1 6 の一部として実行されてもよく、図 1 のTPMS 1 2 8 によって実行されてもよい。

10

20

30

40

50

## 【 0 1 4 7 】

図7に描かれるように、処理は、一組のマイクロタスク生成物の受領に際して開始されてもよい(ステップ702)。たとえば、図1に描かれたMMS 104は、一組のマイクロタスク生成物を分配システム106から受領してもよく、それらのマイクロタスク生成物は処理のためにTPMS 128に転送されてもよい。702において受領された各マイクロタスク生成物は、その対応するマイクロタスクおよび決定された関連付けられた組み合わせられたセグメントにマッピングされる(ステップ704)。マイクロタスクは組み合わせられたセグメントに関連付けられているので、704における処理は本質的には、702において受領された各マイクロタスク生成物を組み合わせられたセグメントにマッピングしている。ある実施形態では、マイクロタスク生成物をマイクロタスクおよびその対応する組み合わせられたセグメントにマッピングするために、マイクロタスク情報146が使用されてもよい。次いで、704において決定された各組み合わせられたセグメントについて、その組み合わせられたセグメントにマッピングするマイクロタスク生成物に基づいて、作業生成物が構築される(ステップ706)。704において決定された組み合わせられたセグメントは次いでその対応するセグメントにマッピングされる(ステップ708)。ある実施形態では、組み合わせられたセグメントをセグメントにマッピングするために、組み合わせられたセグメントおよびその対応するセグメントに関する情報を記憶する組み合わせ情報142が使用される。次いで、708において決定された各セグメントについて、そのセグメントにマッピングする組み合わせられたセグメントについて706において構築された作業生成物に基づいて、作業生成物が構築される(ステップ710)。次いで、706において決定されたセグメントは、個々の入力文書にマッピングされる(ステップ712)。ある実施形態では、これは、セグメント分割情報134を使って実行される。次いで、712において決定された各入力文書について、その入力文書に対応する前記一つまたは複数のセグメントについて710において構築された作業生成物に基づいて、作業生成物が構築される。714において構築された作業生成物はそのタスクについての最終作業生成物を表す。714において構築されたそのタスクについての最終作業生成物に対して、任意的に一つまたは複数の動作が実行されてもよい(ステップ716)。716において実行される動作は、たとえば、最終作業生成物を記憶する、最終作業生成物をタスク要求者に通信する、などを含んでいてもよい。

10

20

## 【 0 1 4 8 】

本発明のある種の実施形態は、タスクを値付けするための技法を提供する。ある実施形態では、本方法は、実行されるべきタスクについての入力情報を受領し、該入力情報を解析して該入力情報の一つまたは複数の属性を決定することを含む。いくつかの実施形態では、前記一つまたは複数の属性は、テキスト文書中の単語の数、オーディオ/ビデオ・コンテンツの長さ、入力情報の複雑さを含んでいてもよい。本方法はさらに、タスクについての値付けを決定するための一つまたは複数の規則の集合を決定することと、前記入力情報の属性および前記規則の集合に基づいて前記タスクについての価格を決定することを含む。

30

## 【 0 1 4 9 】

タスクの値付け

ひとたびタスクおよび/またはマイクロタスクがたとえばMMS 104によって定義されると、そのタスク/マイクロタスクは、分配サブシステムがそのタスクを作業システムまたはコンピュータ・システムに提供するのに先立って、値付けされてもよい。値付けのコンテキストにおいて使用されるタスクは、MMS 104に関して記述したタスクまたはマイクロタスク、あるいは値付けされる必要のある他の任意のタスクであってもよい。

40

## 【 0 1 5 0 】

本発明の諸実施形態は、タスクおよび/またはマイクロタスクについての価格を決定する方法を提供する。本方法は、それに基づいて/それを使ってタスクが実行される場所の入力情報を受領し、該入力情報に関連付けられたタスク記述を受領することを含む。そ

50

の後、前記タスクおよび/または入力情報に関係する一つまたは複数の規則を使って、そのタスクについて価格が決定される。いくつかの実施形態では、同じタスクが、所望される結果または入力情報の型に基づいて異なる値付けをされることがありうる。

【0151】

図9は、本発明のある実施形態に基づく、タスクについての価格を決定するためのシステム900の簡略化された高レベルのブロック図を描いている。システム900は、値付けサブシステム902、入力前処理器904および結果評価器906を有する。図9に描かれるシステム900は、単に本発明の教示を組み込む実施形態の例であって、請求項に記載される本発明の範囲を限定することを意図したものではない。

【0152】

値付けサブシステム902は、タスクに関係したタスク記述960を受領し、任意的に、該タスクがそれに基づいて/それを使って実行される場所の入力情報950を受領し、任意的に、該タスクに関連付けられた任意の制約条件を受領する。値付けサブシステム902は次いで、タスク記述950および任意的に入力情報950に基づいて、タスクを値付けするために適用されるべき、一つまたは複数の値付け規則970を決定してもよい。値付けサブシステム902は次いで、一つまたは複数の適用可能な規則に基づいてそのタスクについての価格を計算する。

【0153】

いくつかの実施形態では、値付けサブシステム902はメモリ・デバイスを含んでもよい。いくつかの実施形態では、メモリ・デバイスはタスクについて価格を決定するためのプログラミング命令を記憶していてもよい。いくつかの実施形態では、メモリ910は、所与のタスクについて価格を決定するために使われるべきさまざまな値付け規則970をも記憶していてもよい。いくつかの実施形態では、メモリ910は、各作業によるタスク実行に関係する統計的情報のデータベースを有していてもよい。この統計的情報は、価格と品質との間の受け入れ可能なトレードオフを達成するために、タスクに値付けし、タスクを作業者に分配するのに使われてもよい。

【0154】

いくつかの事例では、入力前処理器904が、タスクについて価格を決定するための処理の一部として使われてもよい。ある実施形態では、前処理器904は、タスクについての値付けに影響するために、タスクについて受領された入力情報950を修正するために使われてもよい。たとえば、ある種の事例では、前処理器904は、タスクが値付けされるのに先立ってタスクについての入力情報を処理し、入力情報をそのタスクについて決定される価格を下げうる形に変換するよう構成されていてもよい。たとえば、入力情報がテキスト、グラフィックおよび画像を含む文書のラスタ化画像であり、タスクが文書をデータベースに入力するためのフォーマットに変換することである場合、テキスト、グラフィックおよび画像が個々のセグメントに分離され、個々に値付けされて、タスクの全体的なコストを下げるができる。

【0155】

入力前処理器904はまた、結果評価器906および値付けサブシステム902からの入力をも受け取り、その情報を、入力情報950を修正するために使ってもよい。いくつかの実施形態では、タスクの完了からの結果は、結果評価器960に与えられる。すると、結果評価器906はその結果を、正確さ、完了のための時間およびその他の因子について検査し、その情報を入力前処理器904に与える。その結果に基づいて、入力前処理器904は、全体的な価格と正確さの間のバランスが達成されるよう、入力情報を修正してもよい。たとえば、入力文書の内容がデータベースのデジタル項目に変換されるべきであり、入力文書がタイプされたテキストおよび手書きの図を含むタスクを考える。低価格を維持することが主たる基準であれば、値付けサブシステムは、コンピュータによって実行されるタスクに基づいて価格を計算してもよい。しかしながら、コンピュータがそのタスクを実行し、結果を送り返した後、結果評価器がその結果を正確さについて検証し、テキストは正しく変換されたものの図の変換は非常に低い精度であったことを見出すことがあ

10

20

30

40

50

りうる。この場合、結果評価器はこの情報を入力前処理器に与えてもよい。この情報に  
応答して、入力前処理器は、テキストおよび図を二つの異なるセグメントに分離すること  
によって、元の文書を修正してもよい。テキストを含む第一のセグメントは、変換のタスク  
をコンピュータが実行することに基づいて値付けされ、一方、図を含む第二のセグメント  
は変換のタスクを人間が実行することに基づいて値付けされる。このようにして、人間が  
両方のタスクを実行した場合に比べてコストを低く保ちつつ、結果の精度が改善されても  
よい。いくつかの実施形態では、入力前処理器 906 は上記のマイクロタスク管理システ  
ムの一部であってもよい。

#### 【0156】

結果評価器 906 は、タスクの完了後の結果を受領し、その結果を一つまたは複数の基  
準と突き合わせて測定するよう構成されていてもよい。いくつかの実施形態では、結果評  
価器 906 は、値付けサブシステム 902 からタスクについての価格情報を受領して、価  
格がそのタスクについて指定されている最大価格に一致するかどうかについてのフィード  
バックを与えてもよい。いくつかの実施形態では、顧客が、結果評価器 906 によって受  
領される結果を評価する際に使われるべき一つまたは複数の基準を指定してもよい。いく  
つの実施形態では、結果評価器 906 は、タスクを実行する作業者の品質管理のために  
使用されてもよい。結果評価器 906 による評価に基づいて、各作業者についての品質履  
歴がメモリ 910 に記憶されてもよい。いくつかの実施形態では、品質情報は、特定のタ  
スクについての価格を決定するために、値付けサブシステム 902 によって使用されても  
よい。

#### 【0157】

値付けサブシステム 902 および結果評価器 906 は図 9 では別個のユニットとして図  
示されているが、いくつかの実施形態では、両者は同じ値付けサブシステム 902 の一部  
であってもよい。他の実施形態では、値付けサブシステム 902、結果評価器 906 およ  
び前処理器 904 はより大きなタスク生成および分配システム、たとえば図 1 のマイクロ  
タスク管理システム 100 の相異なるコンポーネントとして実装されてもよい。

#### 【0158】

上記のように、値付けサブシステム 902 は、入力情報に加えて他の入力を受領する。  
値付けサブシステム 902 への入力の一つは、値付け規則 970 である。値付け規則 97  
0 は、入力情報の属性および/または一つまたは複数の変数に基づいていてもよい。たと  
えば、規則は「入力情報がオーディオを含み、タスクがそのオーディオをテキストに転写  
することであれば、タスクは人間の作業者によって実行される」というものであってもよ  
い。このように、値付けサブシステムがオーディオ入力情報についての価格を決定する  
とき、それは自動的に、人間の作業者がタスクを実行することに基づいてタスクを値付け  
することを知らず。いくつかの実施形態では、値付け規則 970 は、値付けサブシステム 90  
2 にハード・コードされていてもよい。他の実施形態では、値付け規則 970 は動的であ  
り、顧客が構成設定可能であってもよい。いくつかの実施形態では、いくつかの値付け規  
則がデフォルト規則として指定されていてもよく、ある型の入力情報に関連付けられてい  
てもよい。

#### 【0159】

いくつかの実施形態では、値付け規則は、入力情報の属性に基づいていてもよい。入力  
情報の属性は、入力情報の型、入力情報の内容、入力情報の複雑さまたは入力情報のコン  
テキストを含んでいてもよい。もちろん、当業者は、これが網羅的なリストではなく、入  
力情報についてのもっと多くの属性が可能であることを認識するであろう。各属性には一  
つまたは複数の要素が関連付けられていてもよい。たとえば、入力情報の属性型は、テキ  
スト、ラスタ化画像、グラフィック、オーディオ情報またはビデオ情報を含んでいてもよ  
い。いくつかの実施形態では、入力情報の内容は単語、図、公式などであってもよい。上  
記の変数のリストは網羅的ではなく、単に例解の目的のために提供されていることは理解  
しておくものとする。

#### 【0160】



ここで、入力情報の属性に基づいて価格がどのようにして決定されうるかのいくつかの例を与えておく。以下に与えられる例は網羅的ではなく、単に上記の概念を明らかにするために示されることは理解しておくものとする。これらの例は一つの属性のみに基づいて価格を決定することを記述するが、実際には、価格決定は、さまざまな置換における複数の属性の相互作用に関わることを注意しておくべきである。

【0161】

1. 入力情報の型 いくつかの実施形態では、入力情報はオーディオ・ストリームであってもよく、タスクはオーディオ情報をテキストに文字起こしすることであってもよい。そのような事例では、値付けサブシステムは、タスクを、人間が実行することに基づいて値付けしてもよい。伝統的には、コンピュータの音声認識機能は貧弱だからである。さら

10

【0162】

2. 入力情報の内容 いくつかの実施形態では、入力情報はテキストの画像だけを含んでいることがある。そのような事例では、単語の総数が決定されてもよく、値付けは単語毎ベースであってもよい。この状況では、コンピュータにタスク（たとえばOCRテキスト認識）を実行させることは、人間の作業者よりも低い価格につながりうる。他の実施形態では、顧客が、人間の作業者だけがこのタスクをオファーされうると指定することもありうる。そのような事例では、価格は、何らかのデフォルトではなく、顧客が与える制約条件を使って決定されてもよい。

【0163】

3. 入力情報の複雑さ いくつかの実施形態では、入力情報は、たとえば医師の手書きによる処方など、判読するのが難しい形であることがある。そのような筋書きでは、値付けサブシステムは、人間がタスクを実行することに基づいて価格を計算するよう指示されてもよい。コンピュータが意味のある結果を提供しうることはとてもありそうもないからである。他の実施形態では、入力情報はオーディオ情報、テキストおよびグラフィックの組み合わせを含むことがある。そのような事例では、入力情報は三つのセグメントにセグメント分割されてもよく、第一のセグメントがオーディオ情報を含み、第二のセグメントがテキストを含み、第三のセグメントがグラフィックを含む。各セグメントについての各タスクは、次いで、デフォルト規則または顧客固有の規則を使って個々に値付けされてもよい。

20

30

【0164】

いくつかの実施形態では、タスクの値付けは、入力情報中のコンテンツの総量に依存してもよい。いくつかの実施形態では、入力情報は、単語、グラフィック、画像などといった多様な複雑さの複数の項目を含んでいてもよい。この事例では、マイクロタスクを生成する目的のために入力情報を適正に解析し、セグメントを生成するためには、複数のアルゴリズムが必要とされることがある。いくつかの実施形態では、入力情報がラスタ化された画像の形である場合、入力画像中の幅広い範囲のエッジを検出するために、エッジ検出アルゴリズム、たとえばキャニー（Canny）エッジ検出が使用されてもよい。エッジ検出は、画像処理およびコンピュータ・ビジョンにおいて、特徴検出および特徴抽出のために使用される。エッジ検出アルゴリズムは、デジタル画像中で、画像明るさが急に変わるまたは不連続をもつ点を同定できる。他の実施形態では、入力画像中の線数、文字数、色数などを検出できるアルゴリズムも使用できる。さらに、画像の規格化された輝度〔ルミネンス〕ヒストグラム、規格化されたエッジ・ヒストグラム、色ヒストグラムも、ラスタ化画像中の内容の複雑さを決定するために使用できる。

40

【0165】

上記のように、タスクについての価格は、入力情報および一つまたは複数の値付け規則970の組に基づいていてもよい。いくつかの実施形態では、値付け規則970は一つまたは複数の変数に基づいていてもよい。たとえば、どの値付け規則が適用されるかを決定するのに使用されうる変数は、タスクの所望される結果、タスクを実行する作業者の地理的な位置、タスクを実行する作業者に利用可能なリソース、同じ入力情報についての他の

50

タスクまたは任意の顧客固有の規則などを含んでいてもよい。各変数は、いくつかの要素を有していてもよい。たとえば、タスクの所望される結果変数は、所望される精度、所望される完了時間などを含んでいてもよい。

【 0 1 6 6 】

以下では、値付け規則を決定するために使用されうる変数のいくつかの例を述べる。変数のこの一覧は網羅的であることを意図したものではないことを注意しておくべきである。下記の例は、値付け規則がどのように決定されうるかを説明するために与えられる。当業者は、規則を決定するためのさらに多くの変数が可能であることを認識するであろう。

【 0 1 6 7 】

1. タスクの所望される結果 いくつかの実施形態では、値付けはタスクの所望される結果に依存してもよい。タスクの所望される結果は、タスク実行後に得られる結果の所望される精度およびタスクについての所望される完了時間を含んでいてもよい。たとえば、高レベルの精度(90%+)を要求するタスクについての値付けは、平均的な精度レベル(50%)しか要求しないタスクについての価格よりも高くなりうる。「タスクの所望される結果」変数のもう一つの要素は、タスクについての所望される完了時間であってもよい。たとえば、急ぎの仕事はしばしば標準的なリードタイムの仕事よりもしばしば費用が高くなる。いくつかの実施形態では、同じ仕事が複数の作業者に送られることがあり、それらの結果において統計的に有意な一致が達成されるまで、それらの結果が比較される。もう一つの実施形態では、何らかの英語のテキストが中国語に変換されることを考える。顧客が米国において昼間であるときに数時間以内にタスクの完了を望む場合、結果として得られる値付け規則は、完了時間制約条件のため、その仕事は米国内で実行されるべきであると示すであろう。この事例において、価格は、タスクが米国内で実行されることに基づいて計算される。しかしながら、顧客が翌日まで待つ用意がある場合には、同じタスクは、異なる規則を使って、たとえばタスクが中国で実行されることに基づいて、値付けされてもよい。タスクが中国で実行されれば、タスクについてのより低いコストにつながりうる。

【 0 1 6 8 】

いくつかの実施形態では、所与の金銭的制約条件、たとえば目標価格以内で最高の品質を得るために、タスクは、作業者によって課される価格および作業者の品質履歴を考慮に入れることによって、より小さなマイクロタスクに分割されてもよい。たとえば、タスクが英語の文書を中国語に変換することである場合を考える。作業者Aは翻訳のために1語当たり3セントかかり、95%の精度である一方、作業者Bは単語当たり1セントで70%の精度であるとする。両方の作業者が翻訳についてのそのジョブに対して入札し、高レベルの精度が必要とされる場合、顧客は作業者Aからの入札を受け容れてもよい。作業者Aの精度が作業者Bよりもずっと高いからである。他の実施形態では、顧客が費やす用意のある金額が固定されている場合、文書は、セグメント1は翻訳が難しい情報を含み、セグメント2は翻訳が簡単な情報を含むというように、二つのセグメントに分割されてもよい。マイクロタスクがセグメント1および2のそれぞれに割り当てられてもよい。セグメント1は作業者Aに送られてもよく、セグメント2は作業者Bに送られてもよい。これは、両方の作業者にそれぞれ文書の一部のみを翻訳させることにより、総コストを目標価格以内に保ちつつ、高い全体的な翻訳精度を達成する助けとなりうる。さらにもう一つの実施形態では、翻訳タスクは作業者Bに与えられてもよい。翻訳の結果は次いで、必要に応じた検証および訂正のために、作業者Aに与えられてもよい。この実施形態において、作業者Aがそのタスクのために費やす時間は短くなる可能性が高い。ゼロから翻訳をするのではなく、作業者Bによって実行された翻訳を検証するだけだからである。これは、翻訳タスクの全体的なコストを下げるるとともに、かなり高いレベルの結果の精度を維持する助けとなりうる。

【 0 1 6 9 】

いくつかの実施形態では、作業者の選択は、タスクについての目標価格に基づくことができる。ある事例では、タスク要求者がタスクについての目標価格を与える場合、MMSシステムは、作業者の総量のうちから、その目標価格に基づいてそのタスクを実行する資格

のある作業者の部分集合を決定できる。次いで、そのタスクは、作業者のその部分集合に対してのみオファーされることができ。たとえば、タスクについての目標価格が5ドルであるとする。その価格に基づいて、MMSは、5ドルという目標価格は米国内でそのタスクを実行するためにかかる費用より低いので、そのタスクは米国内では完了できないと決定できる。この事例では、MMSは、目標価格でそのタスクを実行できる地理的位置における作業者のみを選択し、それにより米国内の作業者を自動的に考慮から外す。

【0170】

#### 2. タスクを実行する作業者の地理的位置

いくつかの実施形態では、作業者の地理的位置が価格計算において使用されてもよい。賃金水準が低い国の作業者は、賃金水準が高い国の作業者よりも、低い価格で所与のタスクを実行する可能性が高い。いくつかの実施形態では、タスクの性質は、たとえば規制制約のため、そのタスクが発生国内で実行される必要があるというものでありうる。そのような事例では、結果として得られる規則は、そのタスクについての価格の計算を、タスクがある地理的領域内で実行されることに基づくよう制約してもよい。

10

【0171】

3. 作業者に必要とされるスキル いくつかの実施形態では、価格はタスクを実行するために作業者によって必要とされる特定のスキルに依存してもよい。たとえば、手書きの図をAutoCAD(商標)描画に変換するタスクは、よい製図スキルをもつ作業者を必要とすることがありうる。この状況において、タスクについての価格は、タスクを実行するために必要とされる作業者の特殊スキルに依存し、高くなる可能性が高い。

20

【0172】

4. 作業者に利用可能なリソース いくつかの実施形態では、タスクについて決定される価格は、タスクを実行する作業者に利用可能なリソースに依存してもよい。たとえば、タスクが作業者が特殊なハードウェアまたはソフトウェアを使うことを要求する場合、そのタスクの価格は高くなることがある。そのような事例では、タスクについての価格は、タスクをより簡単にする(たとえば、特殊なハードウェアおよび/またはソフトウェアの必要性をなくすまたは特殊なリソースを作業者に提供する)ことによって下げることができる。

【0173】

上で論じたように、タスクについての値付けは、さまざまな値付け規則に依存しうる。いくつかの実施形態では、タスクを値付けするために使われる規則を顧客が指定してもよい。たとえば、顧客は、人間の作業者だけに自分のタスクを実行してもらいたいことがありうる。顧客は、この情報を値付けサブシステムに与え、値付けサブシステムは、その顧客によって指定されたタスクについての価格を計算するときは常にこの規則を使う。いくつかの実施形態では、デフォルト規則が価格決定のために使用されてもよい。たとえば、入力情報内容がテキストのみであり、タスクはテキストを電子フォーマットに変換することである場合、値付けサブシステムは、コンピュータ作業者を選び、それに応じてそのタスクを値付けするようプログラムされてもよい。もちろん、顧客または値付けサブシステム・オペレーターは、タスクの個別的な要求に基づき、いかなるデフォルト設定をもオーバーライドしてもよい。

30

40

【0174】

いくつかの実施形態では、タスクについての値付けは、所与の入力情報に関連付けられたタスクの数に依存してもよい。たとえば、入力情報がテキストおよび画像を含むとする。その入力情報について定義された二つのタスクがあることがある。テキストをMSエクセル・ファイル・フォーマットに変換し、画像をMS Visioフォーマットに変換するといったものである。そのような状況では、各タスクについての値付けは、その入力情報に関連付けられた他方のタスクに依存してもよい。いくつかの実施形態では、タスクについての値付けは、その値付けサブシステムによって値付けされる他の入力情報についての他のタスクに依存してもよい。該他のタスクは前記タスクに関係していてもいなくてもよい。

【0175】

50

いくつかの実施形態では、顧客は、タスクについての目標価格を指定してもよい。そのような事例では、値付けサブシステムは、上記の因子の一つまたは複数に基づいてタスクについての価格を計算してもよい。価格計算の結果は、次いで、目標価格と比較されてもよい。計算された価格が目標価格より低ければ、計算された価格は、適切な作業への分配のために分配サブシステムに与えられる。いくつかの実施形態では、顧客は目標価格についての許容差限界を与えてもよい。たとえば、顧客は、目標価格100ドルとともに、±5%の価格許容差を指定してもよい。この事例では、計算された価格が95ドルから105ドルの間にある限り、その価格は承認される。しかしながら、計算された価格が目標価格より高い場合、その情報は入力前処理器に通信されてもよい。すると、入力前処理器は、セグメント分割サブシステムとの関連で、入力情報および/またはタスク属性を修正して10  
もよい。次いで、目標価格に一致するまたは目標価格より低い価格を得るために、第二の価格を計算するために、修正された入力情報が値付けサブシステムに与えられる。たとえば、入力情報がラスタ化画像を含み、ラスタ化画像が図およびその図を作成した人物の名前を含んでいるとする。タスクは、この図を、人物の名前とともに、MS VISIOで再生成することであるとする。顧客はこのタスクについて50ドルの目標価格を設定しているとする。第一のパスでは、画像全体がタスクの値付けのために単一の単位として呈示される。入力情報は主として図を含むので、値付けサブシステムは、上記の一つまたは複数の規則に基づいて、この図をVISIOで再生成するには人間の作業が必要とされると決定し、したがってそのタスクを75ドルと値付けしてもよい。この計算された価格が目標価格と比較される。この事例では計算された価格のほうが高いので、この情報は入力前処理器  
20  
に通信されてもよい。入力前処理器は、セグメント分割器サブシステムとの関連で、入力情報を解析し、画像中にテキスト・コンポーネント(すなわち、人物の名前)があることを判別する。そこで、入力情報は二つのセグメントに分割される。一方のセグメントはテキスト情報を含み、他方のセグメントは図を含む。テキスト情報セグメントについてのタスクは、コンピュータが転写を実行することに基づいて値付けされ、図セグメントについてのタスクは人間が転写を実行することに基づいて値付けされる。転写の一部が人間の作業よりも安価なコンピュータによって実行されるので、この状況における二つのセグメントについての総価格は50ドルであってもよい。このように、入力情報および/またはタスク属性を修正することによってタスクについてより低い価格を得ることが可能となりうる。  
30

#### 【0176】

いくつかの実施形態では、高い精度が必要とされる場合、同じタスクが複数の作業者に送られてもよい。しかしながら、これはタスクの価格を著しく高めることがある。いくつかの実施形態では、複数の人が同じタスクを実行する代わりに、価格を下げるために、タスクそのものが毎回修正されてもよい。たとえば、第一のパスの間に第一の作業者が非常に複雑な文書を転写することを求められてもよい。第一のパスから得られた結果が第二の作業者に送られてもよく、そのタスクは第一の作業者の作業を検証することである。第二の作業者は単に前のタスクの結果を検証するので、第二のパスについての価格はより安くなる。同様に、各パスからの結果は検証のために再提出されてもよい。前のタスクの結果を検証するための後の各タスクは、毎回必要とされる作業が少なくなるので、より安くな  
40  
っていく。これは、高い精度レベルを達成しつつ、タスクについてのより低い全体的な価格につながる。

#### 【0177】

上記のように、システム900はタスクについての価格を決定するために使われてもよい。図10は、本発明のある実施形態に基づく、タスクについての価格を決定するプロセス1000の流れ図である。プロセス1000は、たとえば図9の値付けサブシステム902によって実行されてもよい。ステップ1002において、値付けサブシステムは入力情報を受領する。いくつかの実施形態では、入力情報は上述したように修正される必要があることがある。そのような状況では、入力情報は、ステップ1002で値付けサブシステムに通信される前に、ステップ1004で前処理される。入力情報に加えて、値付けサ  
50

ブシステムはステップ1006においてタスク記述をも受領する。タスク記述は、入力情報に対してまたは入力情報を使って実行されるべき処理についての詳細を与える。たとえば、タスク記述は、入力情報中の図がMS Visioの描画データに変換されるべきであることを指定していてもよい。ひとたび値付けサブシステムが入力情報およびタスク記述を受領すると、値付けサブシステムはステップ1008において入力情報の属性を決定する。入力情報の属性は、入力情報の内容、入力情報の内容の型、入力情報の複雑さなどを含んでいてもよい。上記のように、入力情報のこれらの属性は、タスクについての価格を計算するために使われてもよい。

#### 【0178】

ステップ1010では、値付けサブシステムは、タスクを値付けするために適用されるべき規則を決定する。いくつかの実施形態では、入力情報の特定の型に適用されるある種のデフォルト規則がある。さらに、上記のように、価格決定の一部として使用されるべきある種の制約を顧客が与えてもよい。いくつかの実施形態では、一つまたは複数の規則が入力情報について適用可能であってもよい。ステップ1012では、値付けサブシステムは、決定された規則を適用し、入力情報の属性を使って、要求されたタスクについての価格を決定する。いくつかの実施形態では、ひとたび価格が決定されると、決定された価格は、タスクを作業者に分配するための分配サブシステムに通信されてもよい。いくつかの実施形態では、値付け情報は、目標価格との比較のために、上記のように結果評価器に通信されてもよい。

#### 【0179】

図10に示される具体的な諸ステップは、本発明のある実施形態に基づく、タスクについての価格を決定するある個別的な方法を提供するものであることを理解しておくべきである。代替的な実施形態によれば、他のステップ・シーケンスが実行されてもよい。たとえば、本発明の代替的な実施形態は、上に概説したステップを異なる順序で実行してもよい。さらに、図10に示される個別的なステップは、該個別的なステップに適切なさまざまなシーケンスで実行されてもよい複数のサブステップを含んでいてもよい。さらに、個別的な用途に依存して、追加的なステップが追加されてもよく、ステップが除去されてもよい。当業者は、多くの変形、修正および代替を認識するであろう。

#### 【0180】

本発明の実施形態は、多様な用途において使用されうる。以下のセクションは、上記のタスク生成および値付け技法を使用しうる用途のいくつかを開示する。しかしながら、ここに記載される用途のリストは網羅的ではなく、他の多くの用途が上記の実施形態を使用しうることは理解しておくべきである。

#### 【実施例1】

#### 【0181】

##### 手書き文字認識

印刷および手書きのテキストの認識のためのコンピュータ・アルゴリズムは改善が続いているが、いまだ、そのようなアルゴリズムが評価される基準は人間の認識である。コンテキストを使うことでは人間のほうが優れており、人間はテキスト中の複数の異なる型の歪みを認識し、調整しうる。そのような事例では、上記のマイクロタスクの生成および値付けのための技法は、伝統的な手書きテキスト認識方法に対する利点を提供しうる。

#### 【0182】

図11は、手書きテキストのタイプされたテキストへの変換を示している。図11は、タイプされたテキスト1104に変換される必要のある手書きのアドレス帳項目1102を示している。いくつかの実施形態では、結果として得られるタイプされたテキスト1104はデータベース中に入力されてもよい。手書きテキストをタイプされたテキストに変換するサービスは、タスク毎ベースで顧客に提供されてもよい。いくつかの実施形態では、サービスは、タスクを提出するために顧客がアクセスしうるポータルを介して実装されてもよい。いくつかの実施形態では、ポータルはワールド・ワイド・ウェブを通じてアクセスされてもよい。ある実施形態では、いくつかのマイクロタスクが手書きテキストの変

換のために生成されてもよい。ある実施形態では、顧客がたとえばカメラを使って手書きテキストのイメージを取り込んでもよい。次いで顧客は、顧客装置、たとえば図1に示されるタスク要求者システム102上のウェブ・ブラウザを使ってウェブサイトにアクセスし、取り込まれたラスタ化画像を、取り込まれた画像に対して実行されるべき所望の動作を詳述するタスク記述(task description)およびタスク配置(task disposition)とともに、たとえば図1のマイクロタスク管理システム100に提供してもよい。たとえば、タスク配置は、ジョブの結果を顧客によってアクセス可能なサーバーに通信するというものであってもよい。あるいはまた、配置は、結果を指定された電子メール・アドレスにメールで送ることを含んでいてもよい。

**【0183】**

ある実施形態では、マイクロタスク管理システムはラスタ化画像およびタスク記述をタスク要求者システムから受領する。マイクロタスク管理システムは、顧客が十分なクレジットをもっているか、または請求可能なアカウントもしくは当該サービスを使う十分なパーミッションを持っているかどうかを確認するために検査してもよい。次いで、マイクロタスク管理システムは、ラスタ化画像を、たとえば図1の分配システム106による分配のためにフォーマットしてもよい。いくつかの実施形態では、ラスタ化画像はいくつかのセグメントに分割されてもよく、上記のさまざまな規則および入力情報属性に基づいて、マイクロタスクが各セグメントに割り当てられてもよい。マイクロタスク管理システムはタスク記述および何らかのアカウント情報を作業システムに送信する。いくつかの実施形態では、作業システムに送信されるアカウント情報は、顧客によって提供されるアカウント情報とは異なってもよい。いくつかの実施形態では、顧客は、作業システムにアカウントをもつことは要求されない。

**【0184】**

マイクロタスク管理システムは、作業システムによって提供されるタスク進捗をモニタリングし続ける。マイクロタスク管理システムはタスクを取り消して再発行したり、タスクについて作業者にオファーされる額を変えたり、またはさらにはタスクを変更して再提出したりしてもよい。マイクロタスク管理システムは、複数のタスク要求者装置からの入力情報を組み合わせて、組み合わせられた入力情報について単一のタスクを割り当ててもよい。いくつかの実施形態では、マイクロタスク管理システムは、一つのタスクを複数のタスクに分割し、次いで作業システムに分配してもよい。

**【0185】**

ひとたびマイクロタスク管理システムが作業システムから返された十分なタスク結果を受け取ったら、マイクロタスク管理システムは、タスク生成物管理サブシステムを介して、それらの結果を顧客のタスク要求者システムに、たとえば該結果を電子メールで送る、あるいは該結果を顧客がアクセス可能な位置に記憶することによって、提供してもよい。

**【0186】**

いくつかの実施形態では、タスク要求者システムはカメラを有し、インターネット・アクセス機能をもつモバイル通信デバイスであってもよい。他の実施形態では、タスク要求者システムは、画像と取り込み、その情報をマイクロタスク管理システムに提供しうるスキャナまたは多機能デバイスであってもよい。いくつかの実施形態では、タスク要求者システムは複数の顧客によって共有されていてもよく、それら複数の顧客は同じクレデンシャルを使ってサービスへのアクセスを得てもよい。入力情報およびタスク記述を提供するために使用される他の型のデバイスは、タブレット、スレート(slate)、音楽プレーヤー、自動車などを含む。いくつかの実施形態では、タスク要求者システムは、顧客が、入力情報をタスク要求者システム上に直接手書きすることによって入力情報を提供しうる顧客へのインターフェースを提供してもよい。そのようなデバイスの例は、タッチ入力、たとえば指または何らかの接触に感応する領域を備えた装置、あるいはスタイラスに感応する領域を備えた装置を含む。この事例では、タスク要求者システムによって受領される入力は、タスク要求者システムにすでに利用可能な画像と組み合わせられ、次いでマイクロ

10

20

30

40

50

タスク管理システムに提供されてもよい。もう一つの実施形態では、タスク要求者は、画像ではなく、「ストローク」情報を提供してもよい。ストローク情報は、タッチ入力上の接触点についての情報を含み、画像としてではなく、点のリストとして保存されうる。たとえば、InkML (<http://www.w3.org/TR/InkML/>) はストロークを記録するためにXMLを使う方法を規定している。このように、ある実施形態では、タスク要求者は、ラスタ化された画像ではなくInkMLをマイクロタスク管理システムに提供しうる。マイクロタスク管理システムはそのストローク情報を、作業による認識のために、作業システムおよび/または自動化されたシステムへの提出のために、画像フォーマットに変換してもよい。

【実施例2】

【0187】

#### レシート認識

購入または返品されるすべての品目について、何らかの形のレシート/領収証が購入者に対して発行される。各売り手は、レシートについて独自の書式をもつことがある。しかしながら、ある種の情報はあらゆるレシートに共通である。たとえば、総額および品目情報である。人は、そのようなレシートを大量に、非常に短い期間の間に集めることがありうる。たとえば、三日間の出張に行った人は、さまざまな売り手から100を超えるレシートを受け取ることがありうる。そのようなレシートを追跡・管理することは、非常に面倒で時間がかかることがあり、レシートがのちに経費返還のために必要とされる場合は特にそうである。レシート中の関係する情報が、その人にとって非常に低コストで電子フォーマットに変換できれば有用であろう。

【0188】

図12は、レシート1202の内容が認識され、電子フォーマット1204、たとえばエクセルのワークシートの経費レポートに変換されうるアプリケーションを示している。顧客は、経費レポートを作成するために認識される必要のある一つまたは複数のレシートを持っていることがありうる。タスク要求者システムは、各レシートのラスタ化画像を得るために使用されてもよい。いくつかの実施形態では、各レシートは、異なる情報および/または異なる書式を含んでいることがある。各レシート中の情報の大きな部分は、経費レポートを作成するために必要とされないことがありうる。たとえば、顧客は、タスク記述が日付、店名および総額情報のみの認識を含むべきであることを指示してもよい。マイクロタスク管理システムは各レシートの画像と、日付、店名および総額のリストを生成するためのタスク記述とを受け取る。次いで、マイクロタスク管理システムは、いくつかのマイクロタスクを生成し、作業サブシステムに提出してもよい。この事例では、作業者は、日付、店名および総額を含む各レシート画像の領域を同定するよう求められてもよい。作業者は、各レシートの画像を呈示するウェブサイトを訪問し、所望される各情報(たとえば、日付、店名および総額)についての画像部分を横断してドラッグすることによって、これを行ってもよい。この事例では、各レシート画像は複数の部分画像に分割されてもよい。ここで、各部分画像は要求される情報(すなわち、日付、店名または総額)の少なくとも一つを含む。各部分画像は、レシート画像によって占有される平面内の座標によって特徴付けられてもよい。各部分画像は、レシート画像内でのその位置を使って同定されてもよい。

【0189】

この処理のためには人間の入力特に貴重である。各レシートは金額としてフォーマットされた複数の数を含むことがありうるからである。したがって、コンピュータは重要な数字、たとえば総額を見きわめることができないことがありうる。さらに、日時はレシート上で種々の仕方でフォーマットされていることがあり、そのためコンピュータにとって日時を正確に判別することは難しくなることがある。人間の作業者が、レシート画像上の要求される情報の座標を指示するとき、指示された部分画像が自動化文字認識システムにさらなる処理のために送られてもよい。この場合、複数の部分画像からの結果が、タスク生成物管理サブシステムによって、顧客のための一つの結果に組み合わされてもよい。結果は、スプレッドシートまたは表の形であってよいし、あるいはさらには何らかの会計

10

20

30

40

50

システムへの直接入力であってもよい。あるいはまた、作業システムはレシートの画像を最初は自動化された文字認識システムに提出し、次いで文字認識システムによって生成された記号テキストを人間の作業者に提供してもよい。人間の作業者が、日付、時刻および総額に対応する記号テキストを選択してもよい。さらにもう一つの実施形態では、人間の作業者は日時、店名および額を画像から、何らの自動認識処理もなしに、直接タイプするよう求められてもよい。

### 【実施例 3】

#### 【0190】

##### 名刺認識

名刺は、ビジネス・パーソンの中で職業情報を交換するための最も一般に使われるツールである。人は、たとえば展示会などの短い期間の間でさえ、かなりの量の名刺を集めることがある。たいていの名刺は紙の書式なので、損傷したりなくなったりしがちである。これらの名刺の情報を、簡単な記憶および検索のために、電子フォーマットに変換することが有用であろう。図 1 3 は、本発明のある実施形態に基づく、名刺の情報を電子フォーマットに変換する方法を示している。

10

#### 【0191】

図 1 3 は、本発明のある実施形態に基づいて、電子フォーマット 1 3 0 4 に変換される名刺 1 3 0 2 の内容を示している。名刺には多様なフォーマットがあるが、共通の型の情報を含んでいる。名刺は典型的には人の名前、会社名、電話番号および電子メール・アドレスを含む。場合によっては、名刺はウェブサイト、肩書き、物理的な住所、ファクス番号、ツイッター・アカウントおよびグラフィック、たとえば会社のロゴのような追加的な情報をも含むことがある。名刺は、たとえばビジネス上の会合において、あるいはいくつかの営業訪問をした後に、しばしばグループをなして収集されることがある。

20

#### 【0192】

ある実施形態では、一つまたは複数の名刺のラスタ化された画像がタスク要求者システムによって取得されてもよい。それらの画像はマイクロタスク管理システムに与えられる。それらの画像に基づいて、マイクロタスク管理システムは、作業者に画像中の情報を特定し、可能性としては該情報を再生成するよう求める一つまたは複数のタスク記述を生成する。いくつかの実施形態では、名刺の画像が単語にセグメント分割され、次いでそれらの単語が異なる地理的位置にいる作業者に送られてもよい。いくつかの実施形態では、変換のためのマイクロタスクを生成するのに先立って、異なる名刺からの単語が互いに交ぜられることができる。セグメント分割された単語は、論理的な、だが名刺が属する人物のプライバシーを保持する仕方ではグループ化されることができる。たとえば、電子メール・アドレスと人物の名前は一緒にグループ化されないことがありうる。もう一つの例では、異なる名刺からの電話/ファクス番号が同じサブセグメントにグループ化されることができる。そのサブセグメントに関連するマイクロタスクを実行する作業者にコンテキストを与えるためである。セグメント分割技法の例は非特許文献 1 に記載されており、その内容はあらゆる目的のためにその全体においてここに参照によって組み込まれる。

30

#### 【0193】

名刺のロゴはロゴ検出アルゴリズムを使って抽出されてもよい。いくつかの実施形態では、特定の名刺または名刺の集合について生成されたマイクロタスクは、人間および自動化処理の組み合わせを使って実行されてもよい。作業生成物管理サブシステムはマイクロタスクの結果を、vCards (連絡先情報を維持するための規格)、表、スプレッドシートまたは連絡先格納フォーマットの形で、受け取ってもよい。マイクロタスク管理システムは、その情報を、バックエンド・システム、たとえば会社のCRMシステム、個人の連絡先記憶またはソーシャル・ネットワーキング・ウェブサイトに入力してもよい。結果が直接、ソーシャル・ネットワーキング・システム、たとえばリンクトイン (LinkedIn) に入れられる場合において、マイクロタスク管理システムがソーシャル・ネットワーキング・システムについての顧客クレデンシャルを与えられれば、マイクロタスク管理システムは、名刺からの電子メール・アドレスを使ってソーシャル・ネットワーキング・システムに

40

50



おける接続を要求しうる。図 8 A および 8 B ならびに関連する記述は、本発明のある実施形態を使って実装される名刺認識プロセスの詳細を記述している。

【 0 1 9 4 】

図 8 A および 8 B ならびに関連する記述は、一つまたは複数の名刺から情報を認識するためのマイクロタスクがどのようにして生成されるかも記述している。ある実施形態では、名刺から抽出された複数の単語は、異なる id および / または異なる地理的位置をもつ作業者に送られることができる。ある名刺からの単語が異なる名刺からの単語と混合されてもよい。名刺の企業ロゴが抽出され、自動ロゴ検出アルゴリズムに送られてもよい。セグメント分割された単語は、論理的なセグメントに分類されることができる。ここで、電子メール・アドレスおよび人物の名前が名刺から同定され、プライバシーを保持するために、異なる組み合わせられたセグメントにグループ化される（すなわち、同じセグメントにはグループ化されない）。ある実施形態では、異なる名刺から抽出された電話 / ファクス番号が同じ組み合わせられたセグメントにグループ化されてもよい。作業者にタスクにおけるコンテキスト（数字のコンテキスト）を与えることによって高品質出力の可能性を最大にするためである。

【実施例 4】

【 0 1 9 5 】

#### 図の認識

しばしば、会議の間、人はアイデアを視覚的なフォーマット、すなわち手書きのスケッチで提示するためにホワイトボードを使う。そのような情報を取り込む一つの方法は、会議に参加している誰かがそのスケッチをコピーして、関係人員への分配のために、それを電子フォーマット、たとえばパワーポイント・スライドで再生成することである。しかしながら、そのようなタスクはしばしば関係する人物の貴重な時間をくってしまう。そのようなタスクを、上で論じた本発明の諸実施形態を使ってアウトソーシングすることが有用でありうる。

【 0 1 9 6 】

ある実施形態では、タスク要求者システムによって提供されたラスタ化画像が図または簡単なグラフィックを含むことがありうる。ラスタ化画像中の（線、ボックスのような）さまざまなグラフィック要素の位置は、入力ラスタ化画像に比べて異なることがありうる。このため、入力画像と出力結果を位置合わせし、グラフィックの初期レイアウトを保持することが難しくなることがある。この場合、マイクロタスクまたはいくつかのマイクロタスクが作業者に、図またはグラフィックを、特定のソフトウェア、たとえばパワーポイントまたは Visio を使った電子的な形に変換するよう指示してもよい。図 1 4 A および 1 4 B は、描画コンポーネント、たとえば線、長方形などを含む手書きスケッチ 1 4 0 2 およびテキストが、上記のマイクロタスク生成および値付けシステムを使って所望される電子フォーマットのスケッチ 1 4 1 2 に変換される実施形態を示している。ある実施形態では、スケッチ 1 4 0 2 は、会議の際に普通に使われるホワイトボード上に描かれてもよい。

【 0 1 9 7 】

スケッチ 1 4 0 2 が完了されたのち、スケッチの画像がカメラのような通常の画像取り込み装置の任意のものを使って取り込まれることができる。ひとたび画像が取り込まれると、該画像は上述したマイクロタスク生成システムに入力として提出されてもよい。ひとたびスケッチ 1 4 0 2 の画像がマイクロタスク生成器システムによって受領されると、画像が解析されて、スケッチのテキスト部分およびグラフィック / 描画部分が決定される。ある実施形態では、システムは単語境界を認識してマークを付ける。たとえば、画像で黒で塗りつぶしてグラフィック / 描画部分 1 4 0 4 だけを残す。黒で塗りつぶされたセクションのそれぞれは、追跡目的のために番号付けされる。番号付けは、のちに、もとのスケッチを再構成するときに使われてもよい。次いでシステムはスケッチ 1 4 0 2 に存在する一つまたは複数の単語 1 4 0 6 を解析し、これを画像から分離する。次いで、システムは、塗りつぶされたセクションと単語との間の相関情報を生成し、該相関情報をデータベー

10

20

30

40

50

スに記憶する。

【0198】

マイクロタスク管理システムは次いで二つのマイクロタスクを生成する。スケッチの第一の部分1404を電子フォーマットに変換するための第一のマイクロタスクと、一つまたは複数の単語1406を所望されるフォーマットに変換するための第二のマイクロタスクである。その後、第一および第二のマイクロタスクは、上記した所望される結果および規則に基づいて値付けされることができる。完了すると、第一のマイクロタスクは、もとのスケッチ1402において単語に対応する番号付けされたセグメントを含む描画1408を生じてよい。第二のマイクロタスクは、それぞれ描画1408における番号付けされた各セグメントに対応する単語のリストを生じてよい。いくつかの実施形態では、第二のマイクロタスクは、所望される結果に依存して、コンピュータに送られてもよく、あるいは人間によって実行されてもよい。ひとたびマイクロタスク管理システムがマイクロタスクの結果を受領すると、マイクロタスク管理システムは、たとえば図1のタスク生成物管理サブシステム128を使って、二つの結果を組み合わせ、描画1408における番号セグメントを単語のリスト1410からの対応する単語で置き換えることによって、最終的な描画1412を生成することができる。

10

【0199】

いくつかの実施形態では、スケッチ中の単語を変換するためのマイクロタスクを生成する代わりに、単語1406が自動化文字認識エンジン、たとえばwww.abbyy.comに、解析および変換のために送られてもよい。

20

【0200】

図14Aおよび14Bにおいて与えられる例は単に例解目的のためであって、実施形態を手書きのスケッチ/グラフィックに限定すると解釈されるべきではないことを注意しておく。

【実施例5】

【0201】

チェックボックス認識

マイクロタスクを生成および値付けする方法は、書式データをデータ・マイニングのために解析されるデータに変換するために効果的に使用される。ユーザーが書き込むことを求められるさまざまな形の書式がある。書式は、登録書式、アンケート書式、フィードバック書式などを含みうる。これらの書式の多くは、ユーザーが書き込むことが期待される何らかの形のチェックボックスを含む。しばしば、チェックボックスをどのように埋めるかについては定まった規則がない。ユーザーはしばしば、チェックボックスを埋めるために幅広い範囲のインジケータを与える。いくつかの実施形態では、インジケータは「x」マーク、チェック・マーク、完全に塗りつぶされたチェックボックスなどを含みうる。チェックボックス状態を同定するための多くの自動化されたプロセスは、明瞭な「x」マークや完全に塗りつぶされたボックスは簡単に検出するが、自動化システムの大半は、チェックボックスが部分的に塗りつぶされていたり、チェックボックス内にでたらめなマークが置かれていたりする場合には、チェックボックスの状態を適正に決定することができない。さらに、ユーザーがあるチェックボックスを線で消して別のチェックボックスを選択したりチェックボックスに注釈を付けたりする場合、自動化システムはユーザーの意図を適正に検出しないことがあり、チェックボックスを無効にしてしまうことがある。これらの事例において、人間の作業者はより正確な結果を提供できることがありうる。

30

40

【0202】

ある実施形態では、チェックボックスの画像が作業者に与えられ、作業者がボックスがチェックされているか否かを指示してもよい。次いで、結果が、顧客への送達のために、タスク生成物管理サブシステムに与えられる。いくつかの実施形態では、作業者が同じユーザーからの複数のチェックボックスを呈示される。ユーザーは一貫した仕方でチェックボックスを埋めている可能性が高いからである。これは、チェックボックス状態決定を実行する作業者の精度およびスピードを向上させうる。いくつかの実施形態では、マイクロ

50

タスク管理システムは、マークされていると考えられるチェックボックスおよび空であると考えられるチェックボックスを自動的にグループ化してもよい。この事例では、作業者は、誤って分類されたチェックボックスがあればそれを同定することが求められてもよい。これは、すべてのボックスの状態を指示するよりも迅速になされうる。

【実施例 6】

【0203】

上記のマイクロタスク生成および値付け技法から裨益しうるその他の用途のいくつかは、a) 膨大な量のデータから特定の情報をみつける、b) 文書中の誤りを直す、c) 一致物を見つけるためにさまざまなデータ・セットを比較する、d) 誰かのために方向を決定する、e) 顧客が与えた基準を使ってデータをグループ化する、f) 所与のデータから固有名を抽出する、g) 単語 / 言葉を指定された言語およびフォーマットに翻訳 / 変換する、h) 音声認識および文字起こしならびに i) 所与のデータからのロゴの検出を含む。当業者は、ここで具体的に挙げられていない他の多くの用途が上記の技術を使って実装されうることを認識するであろう。

【0204】

図 15 は、本発明の実施形態を実施するために使用されうるコンピュータ・システム 1500 の簡略化されたブロック図である。さまざまな実施形態において、コンピュータ・システム 1500 は図 1 に示され、上述したシステムのいずれかを実装するために使用されてもよい。たとえば、コンピュータ・システム 1500 は、タスク要求者システム 102、MMS 104、分配システム 106 またはプロバイダー・システムを実装するために使われてもよい。図 15 に示されるように、コンピュータ・システム 1500 は、バス・サブシステム 1504 を介していくつかの周辺サブシステムと通信するプロセッサ 1502 を含む。これらの周辺サブシステムは、メモリ・サブシステム 1508 およびファイル記憶サブシステム 1510 を含む記憶サブシステム 1506 と、ユーザー・インターフェース入力装置 1512 と、ユーザー・インターフェース出力装置 1514 と、ネットワーク・インターフェース・サブシステム 1516 とを含んでいてもよい。

【0205】

バス・サブシステム 1504 は、コンピュータ・システム 1500 のさまざまなコンポーネントおよびサブシステムが意図されたように互いと通信することを可能にするための機構を提供する。バス・サブシステム 1504 は概略的に単一のバスとして示されているが、バス・サブシステムの代替的な実施形態は複数のバスを利用してもよい。

【0206】

ネットワーク・インターフェース・サブシステム 1516 は、他のコンピュータ・システムおよびネットワークへのインターフェースを提供する。ネットワーク・インターフェース・サブシステム 1516 は、他のコンピュータ・システムからデータを受信し、コンピュータ・システム 1500 から他のシステムにデータを送信するためのインターフェースとしてはたらく。たとえば、ネットワーク・インターフェース・サブシステム 1516 は、ユーザー・コンピュータがインターネットに接続できるようにし、インターネットを使った通信を容易にしてもよい。

【0207】

ユーザー・インターフェース出力装置 1512 は、キーボード、マウス、トラックボール、タッチパッドまたはグラフィック・タブレットのようなポインティング・デバイス、スキャナ、バーコード・スキャナ、ディスプレイに組み込まれたタッチスクリーン、音声認識システムのようなオーディオ入力装置、マイクロホンおよび他の型の入力装置を含んでいてもよい。一般に、用語「入力装置」の使用は、コンピュータ・システム 1500 に情報を入力するための可能なあらゆる型の装置および機構を含むことが意図されている。

【0208】

ユーザー・インターフェース出力装置 1514 は、ディスプレイ・サブシステム、プリンタ、ファクス機またはオーディオ出力装置のような非視覚的ディスプレイなどを含んで

10

20

30

40

50

いてもよい。ディスプレイ・サブシステムは、陰極線管（CRT）、液晶ディスプレイ（LCD）のようなフラット・パネル・デバイスまたは投影〔プロジェクション〕デバイスをであってもよい。一般に、用語「出力装置」の使用は、コンピュータ・システム1500から情報を出力するための可能なあらゆる型の装置および機構を含むことが意図されている。

#### 【0209】

記憶サブシステム1506は、本発明の機能を提供する基本的なプログラミングおよびデータ構造体を記憶するコンピュータ可読記憶媒体を提供する。プロセッサによって実行されたときに本発明の機能を提供するソフトウェア（プログラム、コード・モジュール、命令）が記憶サブシステム1506に記憶されていてもよい。これらのソフトウェア・モジュールまたは命令は、プロセッサ（単数または複数）1502によって実行されてもよい。記憶サブシステム1506はまた、本発明に従って使用されるデータを記憶するための貯蔵所をも提供してもよい。記憶サブシステム1506は、メモリ・サブシステム1508およびファイル/ディスク記憶サブシステム1510を有していてもよい。

#### 【0210】

メモリ・サブシステム1508は、プログラム実行の間に命令およびデータを記憶するためのメインのランダム・アクセス・メモリ（RAM）1518と、固定された命令が記憶されている読み出し専用メモリ（ROM）1520を含むいくつかのメモリを含んでいてもよい。ファイル記憶サブシステム1510は、プログラムおよびデータ・ファイルについての非一時的な持続的な（不揮発性の）記憶を提供し、ハードディスクドライブ、フロッピー（登録商標）ディスクドライブと関連するリムーバブル・メディア、コンパクトディスク読み出し専用メモリ（CD-ROM）ドライブ、光学式ドライブ、リムーバブル・メディア・カートリッジおよび他の同様の記憶媒体を含んでいてもよい。ファイル記憶サブシステム1510は、タスクについての入力情報、マイクロタスクを実行することから受領される作業生成物、MMS 104によって使われる規則、そのタスクについて生成される最終作業生成物、実行されるべきタスクに関連付けられた因子および制約条件に関する情報（たとえばリスク、品質などに関係した情報）などのような情報を記憶してもよい。

#### 【0211】

コンピュータ・システム1500は、パーソナル・コンピュータ、ポータブル・コンピュータ、ワークステーション、ネットワーク・コンピュータ、メインフレーム、キオスク、サーバーまたは他の任意のデータ処理システムを含むさまざまな型であることができる。コンピュータおよびネットワークの絶えず変わりつつある性質のため、図15に描かれるコンピュータ・システム1500の記述は、コンピュータ・システムの好ましい実施形態を例解する目的のための個別的な例として意図されているに過ぎない。図15に描かれたシステムより多数または少数のコンポーネントをもつ他の多くの構成が可能である。

#### 【0212】

本発明の個別的な実施形態について記載してきたが、さまざまな修正、変更、代替的な構造および等価物も本発明の範囲内に包含される。本発明の実施形態は、ある特定のデータ処理環境内での動作に制約されるものではなく、複数のデータ処理環境内で動作するものも自由である。さらに、本発明の実施形態について具体的な一連のトランザクションおよびステップを使って記述してきたが、これは本発明の実施形態の範囲を限定することを意図したものではない。

#### 【0213】

さらに、本発明の実施形態についてハードウェアおよびソフトウェアの特定の組み合わせを使って記述してきたが、ハードウェアおよびソフトウェアの他の組み合わせも本発明の範囲内であることは認識しておくべきである。本発明の諸実施形態は、ハードウェアのみ、ソフトウェアのみ、あるいは両者の組み合わせを使って実装されてもよい。

#### 【0214】

したがって、明細書および図面は、制約ではなく例解の意味で見なされるべきである。しかしながら、本発明の広義の精神および範囲から外れることなく、それに追加、除去、削除および他の修正および変更がなしうことは明白であろう。

10

20

30

40

50

いくつかの態様を記載しておく。

〔態様 1〕

処理システムによって、実行されるべきタスクに関する入力情報を受領する段階と、  
前記処理システムによって、前記入力情報を解析して前記入力情報または前記タスクの  
一つまたは複数の属性を決定する段階と、

前記処理システムによって、前記タスクについての値付けを決定する一つまたは複数の  
規則を決定する段階と、

前記処理システムによって、前記入力情報の属性および前記規則に基づいて前記タスク  
の価格を決定する段階と、

前記処理システムによって、前記タスクを実行する作業者を決定する段階とを含む、  
方法。

10

〔態様 2〕

態様 1 記載の方法であって、さらに：

前記タスクを実行する複数のサブタスクを決定する段階と、

前記複数のサブタスクからの第一のサブタスクについて：

前記第一のサブタスクに関連付けられた前記入力情報を解析する段階と、

前記第一のサブタスクに関連付けられた前記入力情報および前記タスクの前記価格に  
部分的に基づいて、前記第一のサブタスクの第一の価格を決定する段階とを含む、

方法。

20

〔態様 3〕

態様 1 記載の方法であって、前記タスクが全体タスクについて決定される複数のタスク  
の一部である、方法。

〔態様 4〕

態様 3 記載の方法であって、前記タスクの前記価格が、前記全体タスクの目標価格に基  
づいて決定される、方法。

〔態様 5〕

態様 1 記載の方法であって、さらに：

前記処理システムによって、前記タスクの完了から生じる結果を受領する段階と、

前記処理システムによって、前記結果の精度のレベルが前記タスクについての期待され  
る精度値より低いことを判別する段階と、

30

前記処理システムによって、前記結果の精度のレベルについての情報をセグメント分割  
器サブシステムに通信する段階と、

前記処理システムによって、前記セグメント分割器サブシステムから修正された入力情  
報を受領する段階と、

前記処理システムによって、前記修正された入力情報に基づいて前記タスクの第二の価  
格を決定する段階とを含む、

方法。

〔態様 6〕

態様 5 記載の方法であって、前記修正された入力情報が、前記入力情報からの、コンピ  
ュータによって処理可能な情報を含む第一の部分と、前記入力情報からの、人間によって  
処理可能な情報を含む第二の部分とを含む、方法。

40

〔態様 7〕

メモリおよび該メモリに結合されたプロセッサを有するシステムであって、前記プロセ  
ッサは：

実行されるべきタスクに関する入力情報を受領する段階と、

前記入力情報を解析して前記入力情報の一つまたは複数の属性を決定する段階と、

前記タスクについての値付けを決定する一つまたは複数の規則を決定する段階と、

前記入力情報の属性および前記規則に基づいてタスクの価格を決定する段階と、

前記タスクを実行する作業者を決定する段階とを実行するよう構成されている、  
システム。

50

## 〔態様 8〕

態様 7 記載のシステムであって、前記プロセッサはさらに：  
 前記タスクを実行する複数のサブタスクを決定する段階と、  
 前記複数のサブタスクからの第一のサブタスクについて：  
 前記第一のサブタスクに関連付けられた前記入力情報を解析する段階と、  
 前記第一のサブタスクに関連付けられた前記入力情報および前記タスクの前記価格に部分的に基づいて、前記第一のサブタスクの第一の価格を決定する段階とを執行するよう構成されている、  
 システム。

## 〔態様 9〕

態様 7 記載のシステムであって、前記タスクが全体タスクについて決定される複数のタスクの一部である、システム。

## 〔態様 10〕

態様 9 記載のシステムであって、前記プロセッサがさらに、前記タスクの前記価格を、前記全体タスクの目標価格に基づいて決定するよう構成されている、システム。

## 〔態様 11〕

態様 7 記載のシステムであって、前記プロセッサがさらに：  
 前記タスクの完了から生じる結果を受領する段階と、  
 前記結果の精度のレベルが前記タスクについての期待される精度値より低いことを判別する段階と、  
 前記結果の精度のレベルについての情報をセグメント分割器サブシステムに通信する段階と、  
 前記セグメント分割器サブシステムから修正された入力情報を受領する段階と、  
 前記修正された入力情報に基づいて前記タスクの第二の価格を決定する段階とを執行するよう構成されている、  
 システム。

## 〔態様 12〕

態様 11 記載のシステムであって、前記修正された入力情報が、前記入力情報からの、コンピュータによって処理可能な情報を含む第一の部分と、前記入力情報からの、人間によって処理可能な情報を含む第二の部分とを含む、システム。

## 〔態様 13〕

態様 7 記載のシステムであって、前記プロセッサがさらに：  
 第一のセグメントおよび第二のセグメントを受領する段階であって、前記第一および第二のセグメントのそれぞれは前記入力情報の一部を表し、前記第一のセグメントは前記第二のセグメントとは異なる、段階と、  
 前記規則からの一つまたは複数の規則を使って、前記第一のセグメントの価格および前記第二のセグメントの価格を決定する段階とを執行するよう構成されている、  
 システム。

## 〔態様 14〕

態様 7 記載のシステムであって、前記プロセッサがさらに：  
 前記タスクの前記価格を前記タスクの目標価格と比較する段階と、  
 前記タスクの前記価格が前記目標価格より高い場合、前記入力情報を修正するようセグメント分割器サブシステムに命令する段階と、  
 前記命令にตอบสนองして修正された入力情報を受領する段階と、  
 前記修正された入力情報を使って、実行されるべきタスクの第二の価格を決定する段階とを執行するよう構成されている、  
 システム。

## 〔態様 15〕

態様 14 記載のシステムであって、前記第二の価格が前記目標価格の所定の許容範囲内である、システム。

10

20

30

40

50

## 〔態様 16〕

タスクの価格を決定するようプロセッサを制御する複数の命令を記憶しているコンピュータ可読記憶媒体であって、前記複数の命令は：

前記プロセッサに、実行されるべきタスクに関する入力情報を受領させる命令と、  
前記プロセッサに、前記入力情報を解析して前記入力情報の一つまたは複数の属性を決定させる命令と、

前記プロセッサに、前記タスクについての値付けを決定する規則を決定させる命令と、  
前記プロセッサに、前記入力情報の属性および前記規則に基づいてタスクの価格を決定させる命令と、

前記プロセッサに、前記タスクを実行する作業者を決定させる命令とを含む、  
コンピュータ可読記憶媒体。

10

## 〔態様 17〕

態様 16 記載のコンピュータ可読記憶媒体であって、前記複数の命令はさらに：

前記プロセッサに、前記タスクを実行する複数のサブタスクを決定させる命令と、  
前記複数のサブタスクからの第一のサブタスクについて：

前記プロセッサに、前記第一のサブタスクに関連付けられた前記入力情報を解析させる命令と、

前記プロセッサに、前記第一のサブタスクに関連付けられた前記入力情報および前記タスクの前記価格に部分的に基づいて、前記第一のサブタスクの第一の価格を決定させる命令とを含む、

20

コンピュータ可読記憶媒体。

## 〔態様 18〕

態様 16 記載のコンピュータ可読記憶媒体であって、前記タスクが全体タスクについて決定される複数のタスクの一部である、コンピュータ可読記憶媒体。

## 〔態様 19〕

態様 18 記載のコンピュータ可読記憶媒体であって、前記複数の命令がさらに、前記プロセッサに、前記タスクの前記価格を、前記全体タスクの目標価格に基づいて決定させる命令を含む、コンピュータ可読記憶媒体。

## 〔態様 20〕

態様 16 記載のコンピュータ可読記憶媒体であって、前記複数の命令がさらに：

前記プロセッサに、前記タスクの完了から生じる結果を受領させる命令と、

前記プロセッサに、前記結果の精度のレベルが前記タスクについての期待される精度値より低いことを判別させる命令と、

30

前記プロセッサに、前記結果の精度のレベルについての情報を前記セグメント分割器サブシステムに通信させる命令と、

前記プロセッサに、前記セグメント分割器サブシステムから修正された入力情報を受領させる命令と、

前記プロセッサに、前記修正された入力情報に基づいて前記タスクの第二の価格を決定させる命令とを含む、

コンピュータ可読記憶媒体。

40

## 〔態様 21〕

態様 20 記載のコンピュータ可読記憶媒体であって、前記修正された入力情報が、前記入力情報からの、コンピュータによって処理可能な情報を含む第一の部分と、前記入力情報からの、人間によって処理可能な情報を含む第二の部分とを含む、コンピュータ可読記憶媒体。

## 〔態様 22〕

態様 16 記載のコンピュータ可読記憶媒体であって、前記複数の命令がさらに：

前記プロセッサに第一のセグメントおよび第二のセグメントを受領させる命令であって、前記第一および第二のセグメントのそれぞれは前記入力情報の一部を表し、前記第一のセグメントは前記第二のセグメントとは異なる、命令と、

50

前記プロセッサに、前記規則からの一つまたは複数の規則を使って、前記第一のセグメントの第一の価格および前記第二のセグメントの第二の価格を決定させる命令とを含む、コンピュータ可読記憶媒体。

〔態様 2 3〕

態様 2 2 記載のコンピュータ可読記憶媒体であって、前記第二の価格が前記目標価格の所定の許容範囲内である、コンピュータ可読記憶媒体。

〔態様 2 4〕

態様 2 2 記載のコンピュータ可読記憶媒体であって、前記第一のセグメントの第一の価格および前記第二のセグメントの第二の価格を決定することが：

前記第一のセグメントの第一の価格を、前記規則からの第一の規則を使って決定し、  
前記第二のセグメントの第二の価格を、前記規則からの、前記第一の規則とは異なる第二の規則を使って決定することを含む、  
コンピュータ可読記憶媒体。

10

〔態様 2 5〕

態様 1 6 記載のコンピュータ可読記憶媒体であって、前記入力情報の前記一つまたは複数の属性は、前記入力情報の型、前記入力情報の内容、前記入力情報の複雑さまたは前記入力情報のコンテキストを含む、コンピュータ可読記憶媒体。

〔態様 2 6〕

態様 1 6 記載のコンピュータ可読記憶媒体であって、前記規則からの少なくとも一つの規則は、前記タスクの完了の所望される時間、前記タスクについて所望される精度のレベル、前記タスクを実行する位置、前記タスクを実行するために必要とされるスキル、前記タスクを実行するために提供されるリソースまたは前記タスクの難しさのレベルのうちの一つまたは複数に基づく、コンピュータ可読記憶媒体。

20

〔態様 2 7〕

態様 1 6 記載のコンピュータ可読記憶媒体であって、前記入力情報がテキスト、画像、グラフィック、ビデオ情報またはオーディオ情報のうちの少なくとも一つを含む、態様 1 6 記載のコンピュータ可読記憶媒体。

## 【符号の説明】

### 【 0 2 1 5 】

- 1 0 0 システム
- 1 0 2 タスク要求者システム
- 1 0 4 マイクロタスク管理システム
- 1 0 6 分配システム
- 1 0 8 作業者システム
- 1 1 0 自動化システム
- 1 1 2 タスク要求
- 1 1 6 ユーザー・インターフェース
- 1 1 7 コンテンツ解析サブシステム
- 1 1 8 セグメント分割器サブシステム
- 1 2 0 組み合わせ器サブシステム
- 1 2 2 マイクロタスク生成器サブシステム
- 1 2 4 値付けサブシステム
- 1 2 6 前処理器
- 1 2 8 タスク生成物管理サブシステム
- 1 2 9 アセンブラー
- 1 3 0 タスク規則
- 1 3 2 セグメント分割規則
- 1 3 4 セグメント分割情報
- 1 3 6 セグメント

30

40

50

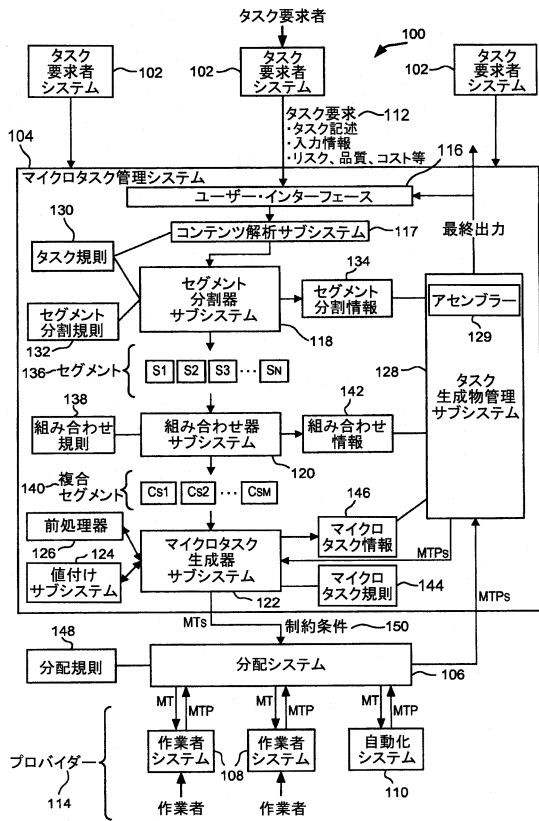


- 1 4 0 複合セグメント ( 組み合わせられたセグメント )
- 1 4 2 組み合わせ情報
- 1 4 4 マイクロタスク規則
- 1 4 6 マイクロタスク情報
- 1 5 0 制約条件
- 1 4 8 分配規則
- 1 1 4 プロバイダー
- 2 0 0 入力情報の内容のプライバシーを保持しつつタスクを実行する方法
- 2 0 2 タスク要求 ( 要求されたタスクのための入力情報を含む ) を受領
- 2 0 3 タスクについての制約条件 ( 例 : リスク、品質等 ) を決定 10
- 2 0 4 入力情報を一組のセグメントに分割。ここで、各セグメントは 2 0 2 で受領された入力情報の内容の部分集合を含む
- 2 0 6 2 0 4 で生成されたセグメントに基づいて一組の複合セグメントを生成
- 2 0 8 2 0 6 で生成された各複合セグメントについて一つまたは複数のマイクロタスクの組を決定
- 2 1 0 2 0 8 で決定された一つまたは複数のマイクロタスクについて値付け情報を決定
- 2 1 1 2 0 8 で決定されたマイクロタスクの組の分配に影響する制約条件を決定
- 2 1 2 マイクロタスクの組を一または複数のプロバイダーに分配 ( 2 1 1 で決定された分配制約条件があればこれに従う )
- 2 1 4 一または複数のプロバイダーから一組のマイクロタスクに対応する作業生成物を受領 20
- 2 1 6 2 1 4 で受領された作業生成物を処理して 2 0 2 で受領 / 決定されたタスクについての最終作業生成物を生成
- 2 1 8 2 1 6 で生成された最終作業生成物に対してアクションを実行
- 3 0 0 入力情報をセグメント分割するために実行される処理
- 3 0 2 タスクについて入力情報を受領
- 3 0 4 タスクについて使われるべき一組のセグメント分割規則を決定
- 3 0 6 3 0 4 で決定された一組のセグメント分割規則を適用して 3 0 2 で受領された入力情報をセグメント分割し、一組のセグメントを生成
- 3 0 8 セグメント分割情報を保存 30
- 3 1 0 3 0 6 で生成された一組のセグメントを組み合わせ器サブシステムに提供
- 4 0 0 複合セグメントを生成するために実行される処理
- 4 0 2 タスクについて生成された一組のセグメントを受領
- 4 0 4 一組のセグメントについて使うべき一組の組み合わせ規則を決定
- 4 0 6 4 0 4 で決定された一組の組み合わせ規則を使って、4 0 2 で受領された一組のセグメントに基づいて一組の複合セグメントを生成
- 4 0 8 組み合わせ情報を保存
- 4 1 0 4 0 6 で生成された一組の複合セグメントをマイクロタスク生成器サブシステムに提供
- 5 0 0 マイクロタスク生成器サブシステムによって実行される処理 40
- 5 0 2 タスクについて生成された一組の複合セグメントを受領
- 5 0 4 5 0 2 で受領された一組の複合セグメントについて使うべき一組のマイクロタスク規則を決定
- 5 0 6 5 0 4 で決定された一組のマイクロタスク規則を使って 5 0 2 で受領された各複合セグメントについて一つまたは複数のマイクロタスクの組を決定
- 5 0 8 5 0 6 で決定された一つまたは複数のマイクロタスクのそれぞれについて値付け情報を決定
- 5 1 0 一組のマイクロタスクについて制約条件を決定
- 5 1 2 5 0 4 で決定された一組のマイクロタスクおよび関連情報を、一または複数のプロバイダーへの分配のために分配システムに提供 50

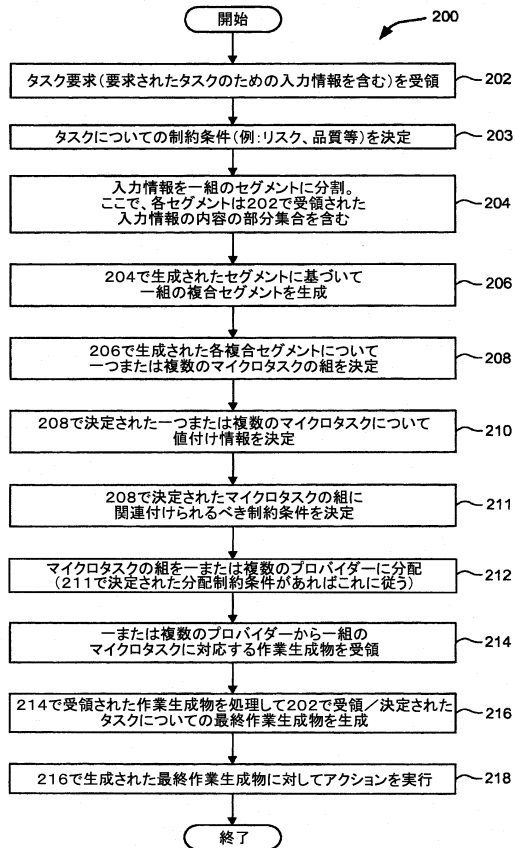
5 1 4	マイクロタスク情報を保存	
6 0 0	分配システムによって実行される処理	
6 0 2	一組のマイクロタスクおよび関連情報を受領	
6 0 4	6 0 2 で受領されたマイクロタスクの組に含まれるマイクロタスクを、該マイクロタスクの組に関連付けられた制約条件が満たされることを保証しつつ、分配	
6 0 6	6 0 4 で分配された一組のマイクロタスクに対応する作業生成物を受領	
6 0 8	6 0 6 で受領された作業生成物をマイクロタスク管理システムに転送	
7 0 0	あるタスクに対応する諸マイクロタスクについて受領されたマイクロタスク生成物に基づいて、該タスクについての最終作業生成物を生成するために実行される処理	
7 0 2	マイクロタスク生成物を受領	10
7 0 4	7 0 2 で受領された各マイクロタスク生成物について、該マイクロタスク生成物に対応するマイクロタスクおよび対応する複合セグメントを決定	
7 0 6	7 0 4 で決定された各複合セグメントについて、該複合セグメントにマッピングされるマイクロタスク生成物に基づいて、該複合セグメントについての作業生成物を構築	
7 0 8	組み合わせ情報を使って、7 0 6 で決定された複合セグメントをその対応するセグメントにマッピング	
7 1 0	7 0 8 で決定された各セグメントについて、そのセグメントに対応する複合セグメントについて7 0 6 で構築された作業生成物に基づいて、そのセグメントについての作業生成物を構築	
7 1 2	セグメント分割情報を使って、7 0 8 で決定されたセグメントに対応する入力文書を決定	20
7 1 4	7 1 2 で決定された各入力文書について、該入力文書に対応する諸セグメントについて7 1 0 で構築された作業生成物に基づいて、該入力文書についての作業生成物を構築	
7 1 6	7 1 4 で生成された最終タスク生成物に対してアクションを実行	
8 0 0	タスク	
8 0 2	入力情報	
8 0 8	セグメント	
8 1 4	セグメント	
8 0 4	スキャンされた画像	30
8 0 6	スキャンされた画像	
8 0 8	セグメント	
8 1 0	画像セグメント	
8 1 2	単語セグメント	
8 1 4	セグメント	
8 1 6	画像セグメント	
8 1 8	単語セグメント	
8 2 0	複合セグメント	
8 2 2	複合セグメント	
8 2 4	複合セグメント	40
8 2 6	複合セグメント	
8 4 0	作業員 1 からのMTP ( マイクロタスク生成物 )	
8 4 2	作業員 2 からのMTP	
8 4 4	作業員 3 からのMTP	
8 4 6	作業員 4 からのMTP	
9 0 0	タスクについての価格を決定するためのシステム	
9 0 2	値付けサブシステム	
9 0 4	入力前処理器	
9 0 6	結果評価器	
9 5 0	入力情報	50

9 6 0	タスク記述	
9 7 0	値付け規則	
1 0 0 0	タスクについての価格を決定するプロセス	
1 0 0 2	入力情報を受領	
1 0 0 4	前処理を実行	
1 0 0 6	タスク記述を受領	
1 0 0 8	入力情報の属性を決定	
1 0 1 0	タスクを値付けするための一つまたは複数の規則を決定	
1 0 1 2	2 1 0 で決定された規則および 2 0 8 で決定された属性に基づいてタスクの価格を決定	10
1 1 0 2	手書きのアドレス帳項目	
1 1 0 4	タイプされたテキスト〔テキスト・データ〕	
1 2 0 2	レシート	
1 2 0 4	電子フォーマット	
1 3 0 2	名刺	
1 3 0 4	電子フォーマット	
1 4 0 2	手書きスケッチ	
1 4 0 4	グラフィック / 描画部分	
1 4 0 6	単語	
1 4 0 8	描画データ	20
1 4 1 0	単語のリスト	
1 5 0 0	コンピュータ・システム	
1 5 0 2	プロセッサ (単数または複数)	
1 5 0 4	バス・サブシステム	
1 5 0 6	記憶サブシステム	
1 5 0 8	メモリ・サブシステム	
1 5 1 0	ファイル記憶サブシステム	
1 5 1 2	ユーザー・インターフェース : 入力装置	
1 5 1 6	ネットワーク・インターフェース	
1 5 1 8	RAM	30
1 5 2 0	ROM	

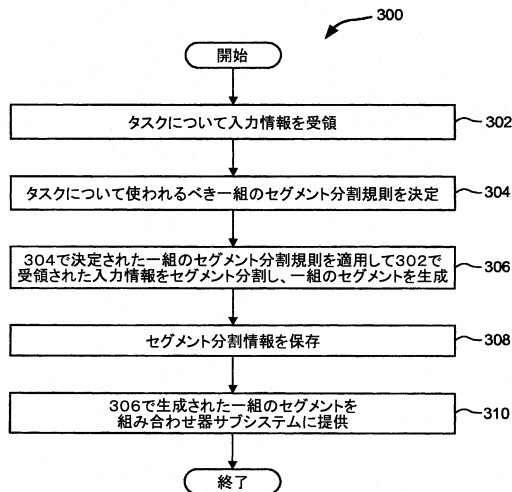
【図1】



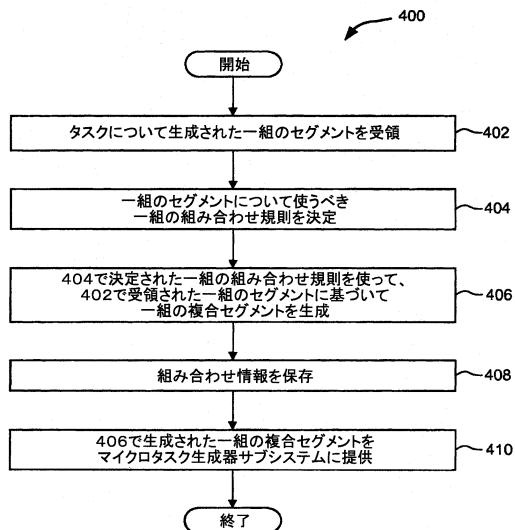
【図2】



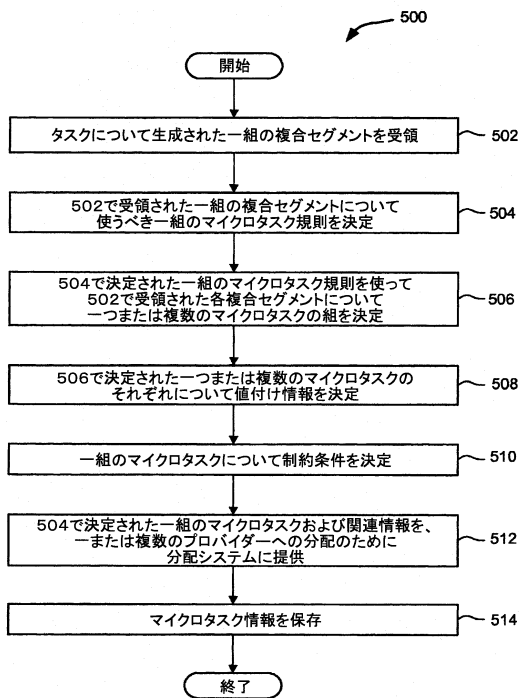
【図3】



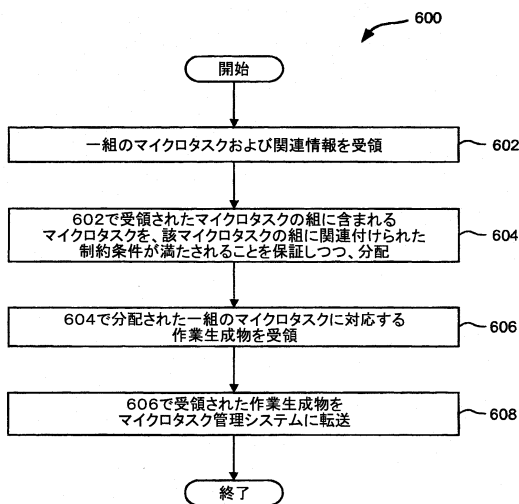
【図4】



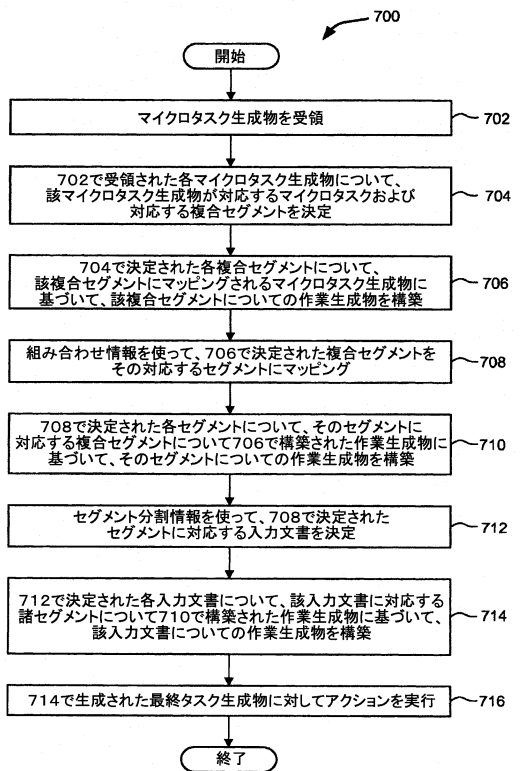
【図5】



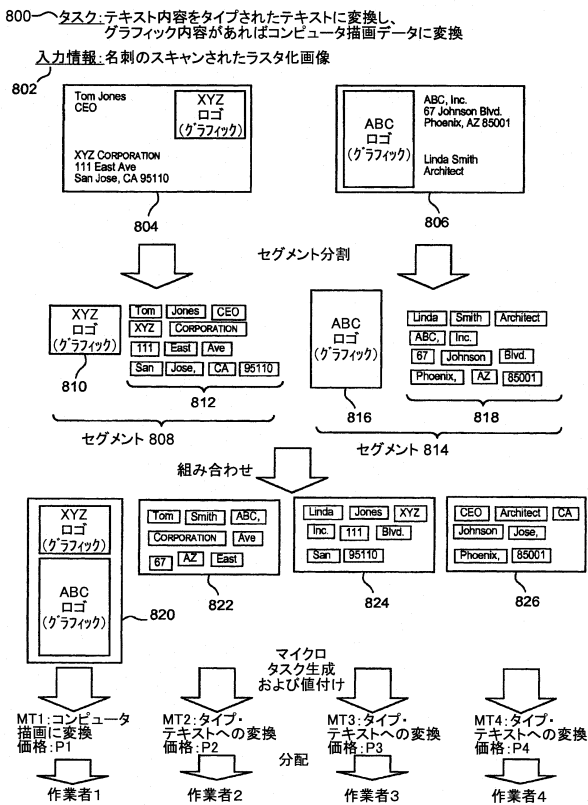
【図6】



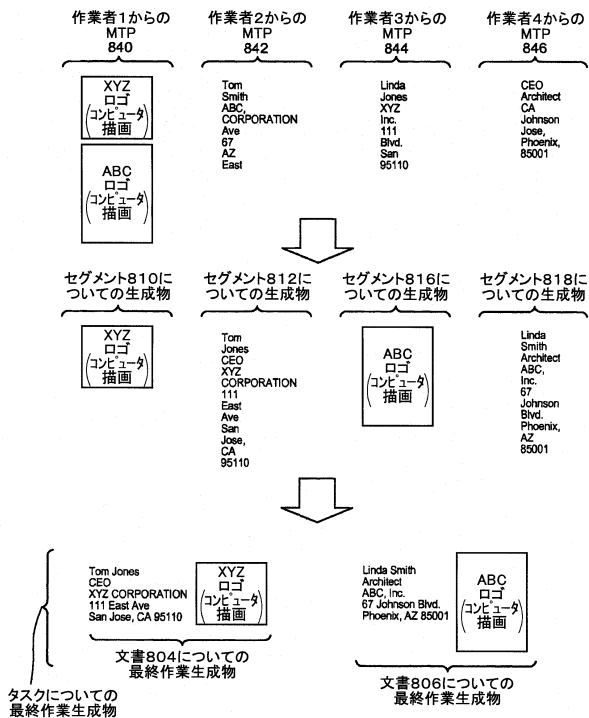
【図7】



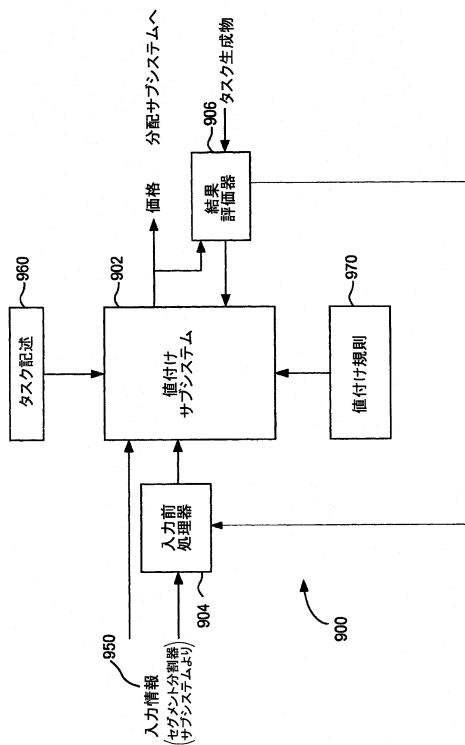
【図8A】



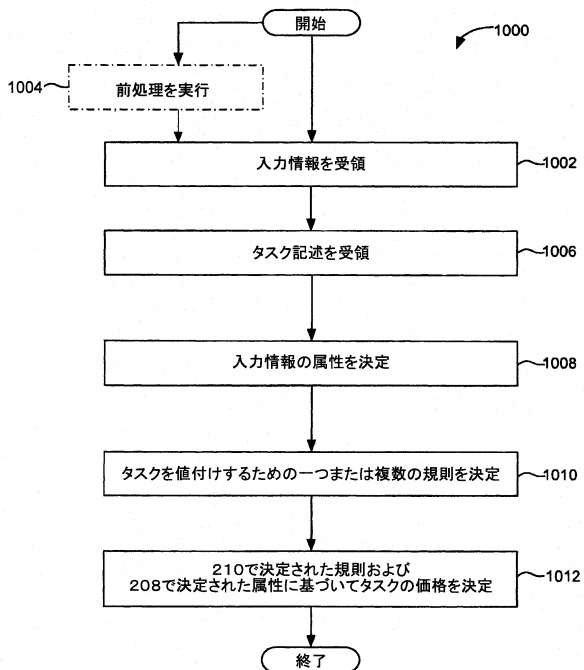
【図8B】



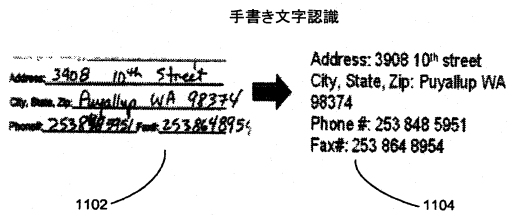
【図9】



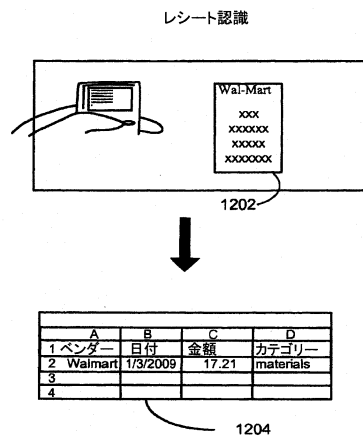
【図10】



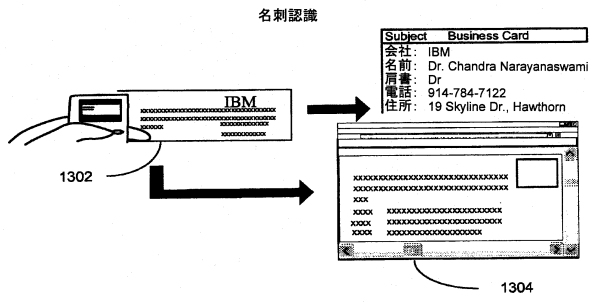
【図11】



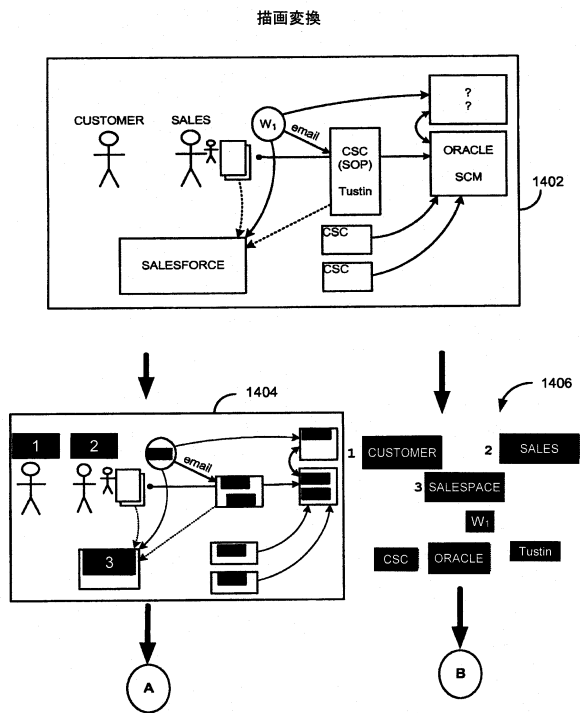
【図12】



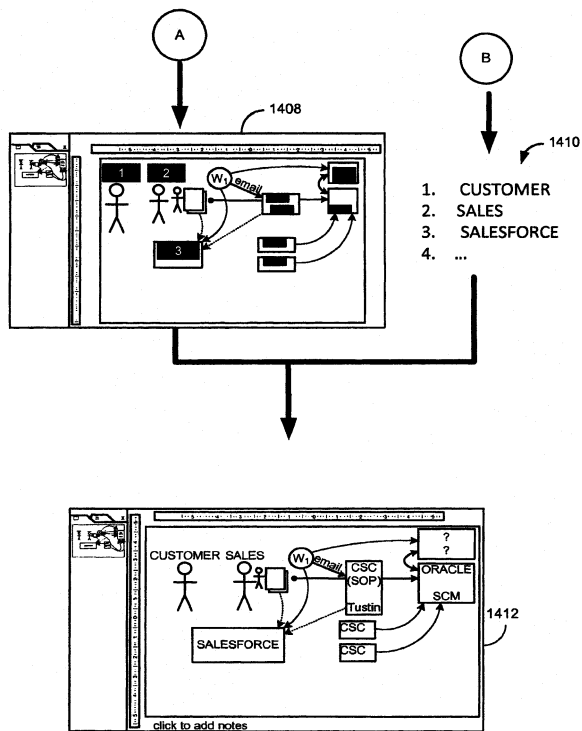
【図13】



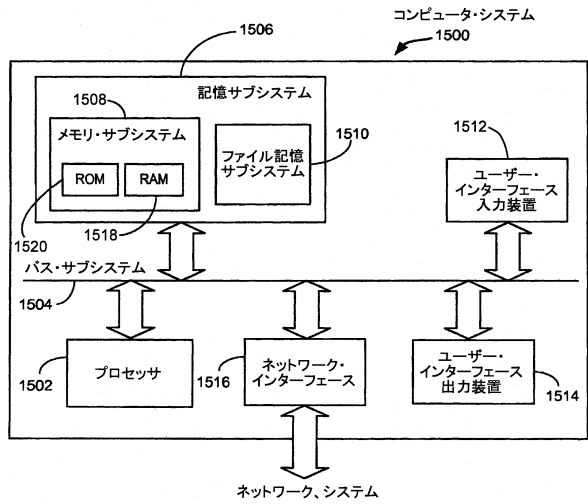
【図14A】



【図14B】



【図15】



## フロントページの続き

- (72)発明者 ジョージ モラレダ  
アメリカ合衆国, カリフォルニア 94025-7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内
- (72)発明者 マイケル ゴーミッシュ  
アメリカ合衆国, カリフォルニア 94025-7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内
- (72)発明者 ティモシー バヨール  
アメリカ合衆国, カリフォルニア 94025-7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内
- (72)発明者 シュイ リウ  
アメリカ合衆国, カリフォルニア 94025-7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内
- (72)発明者 デイヴィッド ジー ストーク  
アメリカ合衆国, カリフォルニア 94025-7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内

審査官 加舎 理紅子

- (56)参考文献 特開2009-223594(JP, A)  
特開2004-185171(JP, A)  
米国特許出願公開第2009/0240549(US, A1)  
米国特許出願公開第2006/0106774(US, A1)  
米国特許第06741737(US, B1)

- (58)調査した分野(Int.Cl., DB名)  
G06Q 10/00 - 50/34