

IBM eDiscovery Analyzer



# ユーザー・ガイド

バージョン 2.1



IBM eDiscovery Analyzer



# ユーザー・ガイド

バージョン 2.1

**お願い**

本書および本書で紹介する製品をご使用になる前に、23 ページの『特記事項』に記載されている情報をお読みください。

本書は、IBM eDiscovery Analyzer のバージョン 2、リリース 1 (製品番号 5724-V36)、および新しい版で明記されていない限り、以降のすべてのリリースおよびモディフィケーションに適用されます。

お客様の環境によっては、資料中の円記号がバックスラッシュと表示されたり、バックスラッシュが円記号と表示されたりする場合があります。

原典： SC19-2707-00  
IBM eDiscovery Analyzer  
User Guide  
Version 2.1

発行： 日本アイ・ビー・エム株式会社

担当： ナショナル・ランゲージ・サポート

第1刷 2008.11

© Copyright International Business Machines Corporation 2008.

---

## 目次

<b>ibm.com および関連リソース</b> . . . . .	<b>v</b>
ご意見の送付方法 . . . . .	v
IBM への連絡 . . . . .	vi
<b>第 1 章 IBM eDiscovery Analyzer の概 要</b> . . . . .	<b>1</b>
検索ストラテジー . . . . .	1
<b>第 2 章 ケースの検索</b> . . . . .	<b>5</b>
検索のヒント . . . . .	6
検索結果の解釈 . . . . .	16

カテゴリーによる検索結果の絞り込み . . . . .	17
検索の保存 . . . . .	18
<b>第 3 章 フラグによる文書の関連性の識別</b> . . . . .	<b>21</b>
特記事項 . . . . .	23
用語集 . . . . .	27
索引 . . . . .	33



---

## ibm.com および関連リソース

製品サポートおよび資料は、ibm.com から入手できます。

### サポートおよびアシスタンス

製品サポートは Web から入手できます。以下の製品 Web サイトで「Support」をクリックします。

#### IBM eDiscovery Analyzer

<http://www.ibm.com/software/data/content-management/support/ediscovery-analyzer/>

#### IBM eDiscovery Manager

<http://www.ibm.com/software/data/content-management/support/ediscovery-manager/>

### インフォメーション・センター

製品資料は、製品のインストール時にインストール可能な Eclipse ベースのインフォメーション・センターで表示できます。デフォルトで、インフォメーション・センターは Web サーバー・モードで実行され、別の Web ブラウザーからアクセスできます。また、ワークステーションでローカルに実行することもできます。インフォメーション・センターは、<http://publib.boulder.ibm.com/infocenter/email/v1r0m0/topic/com.ibm.email.doc/welcome.htm>で参照してください。

### PDF 資料

ご使用のオペレーティング・システムで Adobe Acrobat Reader を使用し、PDF ファイルをオンラインで表示できます。Acrobat Reader がインストールされていない場合は、Adobe の Web サイト (<http://www.adobe.com>) からダウンロードできます。

以下の PDF 資料 Web サイトを参照してください。

製品	Web サイト
IBM eDiscovery Analyzer	<a href="http://www.ibm.com/support/docview.wss?rs=3533&amp;context=SSJKLP&amp;uid=swg27013410">http://www.ibm.com/support/docview.wss?rs=3533&amp;context=SSJKLP&amp;uid=swg27013410</a>
IBM eDiscovery Manager	<a href="http://www-01.ibm.com/support/docview.wss?rs=3506&amp;context=SS8JHU&amp;uid=swg27012725">http://www-01.ibm.com/support/docview.wss?rs=3506&amp;context=SS8JHU&amp;uid=swg27012725</a>

---

## ご意見の送付方法

IBM にお客様のご意見をお寄せください。

ご意見 (コメント) は、[https://www14.software.ibm.com/webapp/iwm/web/signup.do?lang=en\\_US&source=swg-rcf](https://www14.software.ibm.com/webapp/iwm/web/signup.do?lang=en_US&source=swg-rcf)にあるオンライン・リーダー・コメント・フォームを使用してお送りください。

---

## IBM への連絡

米国およびカナダで IBM カスタマー・サービスに連絡する場合は、1-800-IBM-SERV (1-800-426-7378) に電話してください。

ご利用可能なサービスのオプションについては、以下のいずれかにご連絡ください。

- 米国の場合: 1-888-426-4343
- カナダの場合: 1-800-465-9600

IBM への連絡方法について詳しくは、『Contact IBM』の Web サイト (<http://www.ibm.com/contact/us/>) を参照してください。



---

## 第 1 章 IBM eDiscovery Analyzer の概要

IBM® eDiscovery Analyzer によるアーカイブ E メール文書の検索、分析、および検討の方法について本章から説明を始めていきます。

IBM eDiscovery Analyzer は、ディスカバリー要求に該当する E メールおよび添付ファイルの収集に役立ち、さらに検討が必要であると考えられる文書を法律顧問が特定できるようになります。eDiscovery Analyzer は、通常は外部の法的な検討サービスに送られる文書量の削減に役立ち、結果として、そのような検討のコストを削減できます。

eDiscovery Analyzer の検索エンジンでは、Eメールのさまざまな部分 (送信者および受信者のフィールド (TO、FROM、CC、BCC)、件名、本文、および添付ファイル) を検索します。検索エンジンで Eメールの件名行などの一部分を検索するように指定した場合、検索エンジンでは Eメールのその他の部分は検索されません。

ディスカバリー要求に該当する文書を特定するためにケースを検索すると、法的な検討の前に検索をさらに絞ることができます。以下のようにすることが目的です。

- ケース内の Eメール文書の内容の適切な理解
- Eメール文書の検討の優先順位付け
- 追加の検討の必要がない Eメール文書の特定
- ディスカバリー・オーダーに反応すると考えられる Eメール文書の特定およびフラグ付け

追加検討用に収集された一連の文書は、通常、「反応あり」または「第 2 レベルの検討」などの文書フラグによって特定されます。該当する文書がフラグ付けされると、管理者は eDiscovery Manager を使用して追加検討用に文書をパッケージ化できます。

---

### 検索ストラテジー

1 つ以上の検索ストラテジーを使用して、ディスカバリー要求に該当する、または追加の検討が必要な Eメール文書を特定できます。

#### 検索照会の変更

すべての照会は、編集可能な検索照会ボックスに表示されます。検索は保存して再利用が可能であり、各照会を前後に移動して異なる照会による結果の確認、または照会を編集して新規の照会の実行も可能です。

検索照会の構文規則を知っている場合は、テキスト・ボックスに照会を直接入力できます。

#### ノイズの削減

検索結果のセットには、スパム、自動生成の通知、企業の告知など、無関係な文書が含まれている場合もあります。除外したい文書に頻繁に現れる句や送信者を検索

することにより、検討対象の文書の量を削減できます。例えば、自動生成の通知が特定の E メール・アドレスから送信されることがわかっている場合は、そのアドレスを検索し、検索結果を反応なしとしてフラグ付けします。これにより、反応なしとフラグ付けした文書は、後続の検索から除外できます。

頻繁に現れる句および送信者ドメインにより、どの E メールがノイズであるかを推定できます。結果から頻度別に語句をリストするカテゴリーを使用して、それらの句を特定します。

## 照会セットの検討による検索項目の削減

ノイズを削減することに加えて、法律顧問は、相手の法律顧問と相談して、許容できる最少数の結果になる照会を設定できます。ケースの関係者が同意する照会をインポートすることにより、作業を開始するための検索結果のリストを取得できます。各検索を検討することで、結果の数、および戻される内容の範囲を把握できます。カテゴリー表示を使用して、特定の各照会を、ディスカバリー要求を満たす最小限の結果セットに削減します。これらの新規の照会を保存しエクスポートすることで、外部の関係者と相談しさらに改善できます。

## 関連する句の検索

ケース内で最も頻繁に現れる、句、E メール・アドレス、または名前などを検索する場合は、カテゴリー表示を使用して、検索照会でそれらの語句を含めるかまたは除外します。さまざまな種類の句を検索して、関連のある文書に検索結果を絞り込むことにより、ケースの内容をより明らかにできます。

## 検索対象日付の限定

時刻表を使用することにより、検索対象の日付の範囲を限定できます。Eメールの送信が多い、または少ない日付を検討することにより、ケースの内容への理解度が高まります。さまざまな日付範囲の検討を繰り返すことで、理解度がさらに高まります。

## 無関係な語および句の削減

検索結果において繰り返し出現する無関係の句を制限するには、検索で無視されるテキスト (無視するテキスト) を指定します。例えば、一般的な特記事項を無視するテキストに追加します。検索エンジンでは一般的なストップワードを使用しないため、各ケースごとに語または句を指定する必要があります。

## E メール・スレッドの表示

検索結果から E メール・スレッド (会話) を表示できるので、検討対象の E メール文書に出現する内容および会話が理解しやすくなっています。結果リストで「スレッドの表示」および「スレッド別にソート」の各オプションを使用します。

### 関連概念

16 ページの『検索結果の解釈』

### 関連タスク

5 ページの『第 2 章 ケースの検索』

### 関連資料

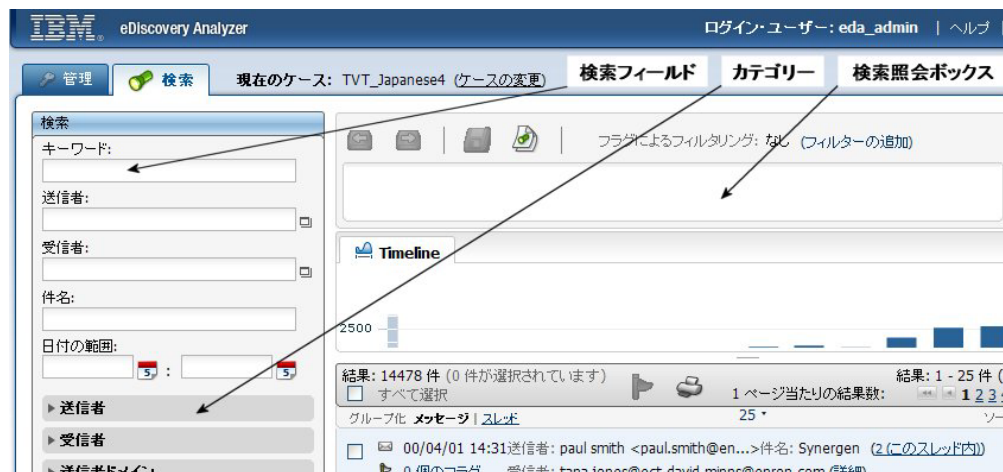
6 ページの『検索のヒント』



## 第 2 章 ケースの検索

ケースを検索して関連のある内容を特定し、ケースのサイズを削減し、さらに検討が必要な文書を特定します。

以下のイメージには、検索インターフェースでの検索フィールド、カテゴリー、および検索照会ボックスの位置が示されています。



ケースを最初に開くと、デフォルトの検索照会が自動的に実行されます。別のデフォルト照会を定義していない限り、照会ではケース内のすべての E メールが戻されます。

ケースを検索するには、以下のようになります。

1. eDiscovery Analyzer にログインし、ケースを選択します。
2. 検索フィールドにキーワード、送信者、受信者、内容、または日付の任意の組み合わせを入力します。少なくとも検索項目を 1 つ入力する必要があります。ケース内のすべての文書を戻すデフォルトの照会は \*:\* であり、すべてのフィールドおよびすべての項目を意味します。
3. 「検索」をクリックします。照会は検索照会ボックスに表示され、結果は検索ページの「結果」および「時刻表」セクションに表示されます。
4. オプション: 検索を絞り込むには、検索フィールドに追加の検索項目を入力して、「検索への追加」を選択し、「検索」をクリックします。新規の検索項目が検索照会のテキスト・ボックスに追加されます。
5. オプション: 別の検索項目も試してみてください。例えば、検索照会ボックスで照会を編集するか、カテゴリーの項目から選択するか、時刻表または「日付の範囲」フィールドを使用して日付範囲を変更するか、または「検索」セクションで条件を追加または削除します。
6. オプション: 「フラグによるフィルタリング」メニューから 1 つ以上のフラグを選択します。これにより、後続の各照会において、選択したフラグが指定された文書を含めたり除外したりできます。「フラグによるフィルタリング」の設定は、ユーザーごとに異なります。

## 関連概念

1 ページの『検索ストラテジー』

## 関連タスク

21 ページの『第 3 章 フラグによる文書の関連性の識別』

---

## 検索のヒント

検索照会演算子を使用して、検索を絞り込むことができます。

すべての検索に対して、以下のルールが適用されます。

- 通常、句読点は単語の区切り記号として扱われます。照会 `case-sensitive` は、完全一致突き合わせ照会 `"case sensitive"` が返すのと同じ文書を返します。ただし、一部のハイフン付きの語 (`self-inflicted` など) は単一の語として扱われます。
- 検索語は、大/小文字の区別がありません。照会 `IBM` では、`ibm` を含む文書も検出されます。
- 検索語は、同じ単語の変化形にも一致します (基本型の一致)。照会 `stock option` は、`stock options` に一致します。照会 `mouse` は `mice` に一致します。また、例えば `king` と入力すると、検索エンジンから語 `king` または `kings` を含むすべての文書が返されます。照会 `king lear` を入力すると、検索エンジンは語 `king` または `kings`、および `lear` を含む文書を返します。

単純キーワード検索を行う場合は、「**キーワード**」フィールドに 1 つ以上の照会用語 (キーワード) を入力します。検索エンジンは、件名、本文、または添付ファイルにこのキーワードのすべてまたはキーワードの変化形が含まれる文書を返します。デフォルトのブール演算子は `AND` です。

さらに正確な結果を得るには、より具体的なキーワードを使用します。例えば、`taxes` ではなく `tax avoidance strategies` で検索します。または、`energy` ではなく `California energy usage` で検索します。

単純キーワード検索で、探しているものではない多くの文書が返される場合は、演算子またはその他の検索語フィールドを使用して検索を絞り込むことができます。

## 用語の除外

用語を除外するには、負符号 (`-`) またはブール値 `NOT` 演算子を使用します。例えば、用語 `lear` が含まれる文書を検索するときに `edward` が含まれる文書は表示しない場合は、照会 `lear -edward` または `lear NOT edward` を入力します。

負符号 (`-`) や `NOT` は、用語だけでなくその変化形にも適用されます。例えば、照会 `-edward` では、語 `edward's` を含む文書も除外されます。

## 用語の完全一致

入力した順序も含め完全に一致する複数の用語を含む文書を検索する場合、二重引用符 (`"`) を使用できます。例えば、`mouse trap` という句を含む文書を表示するが、`trap a mouse` や `mice escaped the trap` などの関連した句に一致する文書は表示しない場合は、照会を二重引用符で囲みます。`mouse` と `mice` など、用語の変化形

は完全一致突き合わせ検索では一致とみなされません。照会 "mouse trap" は mouse traps または mice trap に一致しません。句の中の 2 語の間に用語を入れることはできないため、mouse in trap は一致しません。大/小文字についてはこの検索でも区別されないため、Mouse Trap には一致します。ただし、句読点は無視されるので mouse, trap や See the mouse? Trap it. は一致します。

完全一致突き合わせ検索には、以下のルールが適用されます。

- 用語の順序は、照会内の用語の順序と同一である必要があります。例えば、照会 "life sentence" は sentence to life に一致しません。
- 各用語は同じ形式である必要があります。変化形は一致しません。照会 "stock option" は stock options に一致しません。
- フィールド分けした照会の場合、空白文字で分離した複数の用語は、用語を括弧で囲むことによってグループ化する必要があります。

表 1. 完全一致突き合わせ検索の例

検索対象	検索照会ボックスに入力する照会の例
mouse trap という句 (関連する句を含めない)	"mouse trap"

## 特定の接近性における検索

同じ文、段落、またはある単語範囲の中に現れる用語に結果を制限するには、照会に用語と WITHIN コンテキスト を指定します。指定した用語が指定の順序で現れるようにするには、INORDER を追加します。例えば、照会 WITHIN filed patents in December INORDER を使用すると、filed patents in December に一致するテキストが検出されますが、patents filed in February や patents filed in December には一致しません。

- WITHIN SENTENCE では、同じ文の中に指定した用語を含む文書が一致します。
- WITHIN PARAGRAPH では、同じ段落内に指定した用語を含む文書が一致します。
- WITHIN  $n$  では、それぞれの指定した単語数の中に指定した用語を含む文書が一致します。 $n$  の値は、指定した用語数以上の正数である必要があります。

どのような接近性検索照会でも、文書に一致するためには、接近性検索で指定する用語が文書の同じ部分に現れる必要があります。例えば、接近性照会では、指定した用語が件名フィールドと本文フィールド、または本文と添付ファイルの部分に現れる文書は返されません。

表 2. 接近性検索の例

検索対象	検索照会ボックスに入力する照会の例
同じ段落の中に任意の順序で出現する article but、cancer dancер、elegant funeral の 3 つの句。	("article but" "cancer dancер" "elegant funeral") WITHIN PARAGRAPH
5 つの単語の範囲内に任意の順序で出現する and、begin、court の 3 つの単語 (単語の変化形を含む)。	(and begin court) WITHIN 5
5 つの単語の範囲内にこの順序で存在する and、begin、court の 3 つの単語 (単語の変化形を含まない)。	("and" "begin" "court") WITHIN 5 INORDER



## ブール演算子を使用して検索する

ブール演算子 OR は、照会に指定する用語のうち少なくとも 1 つが含まれる文書が戻されることを指定します。例えば、照会 (othello OR otello) では、用語 othello と otello のいずれかまたは両方を含む文書が返されます。ブール演算子では、大/小文字の区別をしません。

また、括弧を使用してブール演算子 AND、OR、または NOT の組み合わせを使用することもできます。例えば、照会 cougar OR (jaguar AND NOT car) では、cougar を含むすべての文書、および cougar を含まなくても jaguar を含み、かつ car を含まない文書が返されます。

グループ化するには、括弧を使用します。ブール演算子でもある単語を検索するには、その単語を二重引用符で囲むか、前にエスケープ文字 (\) を付ける必要があります。

例:

- 照会 subject:(hedge OR fund) では、hedge と fund のいずれか一方、あるいは hedge または fund の変化形が件名に含まれる文書に一致します。
- AND、OR、NOT などの予約語 (検索において演算子としての意味を持つ単語) を検索するには、これらの単語が二重引用符で囲まれていない場合はエスケープする必要があります。例えば、car AND \not、car AND "not"、"car not" といった照会は作成できますが、car AND NOT という照会は処理されません。

表 3. ブール演算子を使用した検索の例

検索対象	検索照会ボックスに入力する照会の例
件名に存在する単語 hedge、fund、hedging、または funds。	subject:(hedge OR fund)
単語 car および not	car AND \not、または car AND "not"、あるいは "car not"
othello の代替スペリング、または otello、あるいはその両方。	(othello OR otello)

## 不明または形式が変化する文字や用語を検索する

ワイルドカード文字を使用すると、フル・スペルが不明な場合の文書の検索、または用語の変化形を含む文書の検索が容易になります。例えば、照会 czech\* では、用語 czech、czechoslovakia、czechoslovakian、および文字列 czech で始まるその他の単語 (後続文字がない czech も該当) を含む文書が返されます。

照会の先頭にワイルドカード文字を付加すると (例えば、\*zech) 検索エンジンから結果が戻される時間が長くなる場合があります。

また、句検索でワイルドカード文字を使用することもできます。ワイルドカード文字 (\*) は 0 個または複数の文字 (ただし 1 語に限る) に一致します。例えば、照会 "John \* Kennedy" では、John Fitzgerald Kennedy と John F Kennedy が含まれる文書が戻されますが、John Kennedy を含む文書は戻されません。ただし、ワイルドカード文字 (\*) は単語の区切り文字をまたいで一致しないので、"Jo\*ennedy"



は "John Kennedy" など、空白文字を含むその他の変化形には一致しません。ワイルドカード文字 (\*) は、単一の単語内にある、0 個を含む任意の数の文字に一致しますが、アスタリスクは単語の区切り文字 (複数の単語) をまたいで一致しないため、一致するためには少なくとも 1 つの文字を含む単語が存在する必要があります。例えば、"Jo\*dy" は "John Kennedy" (2 語) には一致せず、"John \* Kennedy" は "John Kennedy" (中央の単語がない) には一致しません。あるいは、照会用語 fin\*int は fingerprint には一致しますが、fine print には一致しません。

単一の文字に対するワイルドカード文字には、疑問符 (?) を使用します。検索語 ra?or は、razor や rasor に一致しますが、raptor には一致しません。照会 "John ? Kennedy" では、用語 John F Kennedy、または John F. Kennedy を含む文書が返されますが、John Fitzgerald Kennedy を含む文書は返されません。

表 4. ワイルドカード検索の例

検索対象	検索照会ボックスに入力する照会の例
名前 John Fitzgerald Kennedy または John F Kennedy	person:"John * Kennedy"
名前 John F Kennedy または John F. Kennedy	person:"John ? Kennedy"
名前 Mikhail Gorbachev	person:Mi*1 Gorbachev

## 特殊文字を検索する

" \ ( ) などの特殊文字、またはブール演算子 (AND、OR、NOT) を、検索照会にリテラル用語として含めることができます。これを行うには、特殊文字の前にエスケープ文字 \ を置きます。特殊文字以外の文字の前にエスケープ文字 (\) を置いた場合、エスケープ文字 \ は無視されます。

エスケープ文字が必要な単一文字には、\u3000、\t、\n、\r、+、-、(、)、:、^、@、[、]、\、{、}、~、\*、?、<、>、=、! などの Unicode 文字があります。ただし、照会で二重引用符内に含まれる場合にエスケープ文字が必要な特殊文字は、\*、?、"、\ のみです。

ブール演算子、接近性演算子、算術記号 (<=、>=、!=、) など、検索照会の中にも存在する単語や文字を検索する場合は、その文字や単語の前にバックスラッシュ (\) を挿入します。複数の記号で構成される特殊文字シーケンス (<= や >= など) では、各記号をエスケープ (例えば、\<=) する必要があります。SENTENCE など、アルファベットを含む特殊文字シーケンスの場合は、\SENTENCE のように必要なエスケープ文字は 1 つだけです。

**ヒント:** Lotus Notes® の E メール・アドレスでは、特殊文字 / はエスケープ文字がなくても保存および認識可能です。例えば、"Raul Martinez San Jose" を含む照会では、Lotus Notes® アドレスである Raul Martinez/San Jose/Example が返されます。

表 5. 特殊な文字または単語を使用した検索の例

検索対象	検索照会ボックスに入力する照会の例
より小か等しい記号 (<=)	\<=
単語 car および not	car AND \not

表 5. 特殊な文字または単語を使用した検索の例 (続き)

検索対象	検索照会ボックスに入力する照会の例
括弧付きの句 (but not here)	\(but not here\)

## フィールドまたはカテゴリによる検索

検索フィールドにテキストを入力するか、検索照会ボックスに有効な照会を入力することで、検索を行うことができます。フィールドまたはカテゴリによる検索を行うには、フィールドまたはカテゴリの名前、コロン (:) を入力した後に用語を入力します。例えば、句 "tax evasion" が件名の行に含まれている文書を検出するには、「件名」フィールドに "tax evasion" を入力するか、または検索照会ボックスに `subject:"tax evasion"` を入力します。joe.morales@example.org が送信した E メールを検索するには、「送信者」フィールドに joe.morales@example.org を入力するか、検索照会ボックスに `from:joe.morales@example.org` を入力します。

特定のフィールドまたはカテゴリにのみ一致するよう、照会用語の前にフィールド名またはカテゴリ名を置いて検索を絞り込みます。照会に句の中のない空白スペースがある場合は、照会を括弧で囲んで、句の適切な部分をグループ化します。例えば、`subject:(IBM Software)` では単語 IBM および software が Eメールの件名に含まれている文書だけが返されます。照会 `subject:"IBM Software"` は、IBM Software がこの順序で含まれる文書を返し、これらの単語の変化形は除外されます。

照会用語 `subject:IBM Software` では、IBM が件名に含まれ、software が件名、本文、または添付ファイルに含まれる文書が返されます。あるいは、検索照会 `senderdomains:example.org` では送信側のドメイン名が example.org であるすべての文書が返され、照会 `-senderdomains:example.com` ではドメイン名が example.com ではないすべての文書が返されます。

表 6. カテゴリ検索の例

検索対象	検索照会ボックスに入力する照会の例
句 case sensitive	<code>phrase:"case sensitive"</code>
john.smith@example.com が送信した E メール	<code>senders:"john.smith@example.com"</code>
jill.jones@example.com 宛に送信された E メール	<code>recipients:"jill.jones@example.com"</code>
ドメイン example.com から送信された E メール	<code>senderdomains:"example.com"</code>
ドメイン example.org 宛に送信された E メール	<code>recipientdomains:"domain.example.com"</code>
John Smith という人物に関するすべての記述	<code>person:"John Smith"</code>
ロケーション Silicon Valley に関するすべての記述	<code>location:"silicon valley"</code>
企業 IBM に関するすべての記述	<code>company:IBM</code>
Confidential のフラグが立てられているすべての E メール	<code>flags:"Confidential"</code> フラグ検索は、大/小文字の区別があります。

## E メールを送信者および受信者を検索する

Eメールの送信者や受信者を検索するには、「送信者」または「受信者」検索フィールドに名前または E メール・アドレスを入力するか、「送信者」または「受信者」カテゴリから選択します。このフィールドにアットマーク (@) が含まれるストリングはすべて、E メール・アドレスとみなされます。不完全な E メール・アドレスでは、索引付き Eメールの完全アドレスに一致しません。ワイルドカードを含むアドレスは、不完全とは限りません。例えば、John\*@example では John.Smith@example.com に一致しませんが、John\*@example.com では一致します。あるいは、Smith@example.com は example.com のすべての Smiths に一致しません。アドレスの両方の部分 (ID とドメイン) にワイルドカードを使用できます。例えば、John\*@example\* で John Smith@example.comを検出することができます。「送信者」または「受信者」フィールドの横にあるアイコンをクリックしてテキスト・ボックスを拡大し、Eメール・アドレスの長いリストを入力することが可能です。

また、Eメール・アドレスのフィールド from:、to:、cc:、bcc:、toccbcc: および sentrepresenting: と共に名前や Eメール・アドレスを検索照会ボックスに入力することもできます。これらのフィールドは、それぞれ送信者ヘッダー、to、cc、bcc、toccbcc リスト、またはプリンシパル (on-behalf-of) ヘッダーでのみ検索を行います。

ワイルドカードを使用できます。ワイルドカードは、Eメール・アドレスの単語区切り文字として使用されている句読点に一致しますが、単語の区切り文字をまたいでは一致しません。

Eメール・アドレスの検索は、Eメール・アドレスに存在する可能性がある句読点および特殊文字を考慮して実行されます。Eメール・アドレスでは、大/小文字の区別がありません。1単語でも、Eメール・アドレスまたは送信者や受信者の名前に一致します。用語の代替書式は、Eメール・アドレス検索ではサポートされていません。

例えば、from:"John Doe" は john.doe@example.com からの Eメールには一致しませんが、John Doe <jdoe@example.com> からの Eメールには一致します。キーワード検索または句検索では、単一の名前またはアドレス内でのみマッチングを行います。例えば、照会用語 to:(craig smith) は、Craig Miller および John Smith に宛てられた Eメールに一致します。照会に @ 記号がある場合、照会は Eメール・アドレスの完全一致突き合わせが必要であると解釈されます。例えば、照会 from:smith@example.com は、Eメール・アドレス smith@example.com だけに一致し、john.smith@example.com からの Eメール・アドレスには一致しません。

表 7. Eメール・アドレス検索の例

検索対象	検索照会ボックスに入力する照会の例
John Doe から送信された Eメール	John Doe <jdoe@example.com>
Craig Miller および John Smith に宛てられた Eメール	to:(craig smith)
smith@example.com から送信された Eメール	from:smith@example.com

表 7. E メール・アドレス検索の例 (続き)

検索対象	検索照会ボックスに入力する照会の例
John Smith が送信した E メール	from:"john smith",from:jsm?th@example.com,
John Smith宛での E メール	to:john.sm*th@example.com
Jill Jones にブラインド・コピーが送信された E メール	bcc:jill.jo*es@example.net
John Smith または Jill Jones にコピーが送信された E メール	cc:"john smith" OR cc:"jill jones"
Jill Jones が別の送信者の代理人として送信した E メール	sentrepresenting:"Jill Jones"

## 日付や日付範囲で検索する

予約語 `docdate` を検索照会で使用して、ある日付範囲内に送信された文書に検索を絞り込むことができます。検索照会 `docdate < "2007-12-18"` では、2007 年 12 月 18 日より前に送信された文書だけが返されます。検索照会 `docdate > "2007-12-18"` では、2007 年 12 月 18 日より後に送信された文書だけが返されます。検索照会 `docdate >= "2006-05-16" AND docdate <= "2007-12-18"` では、2006 年 5 月 16 日以降、2007 年 12 月 18 日までの間に送信された文書だけが返されます。

予約語 `postedtime` (E メールが送られた日付) または `deliveredtime` (E メールが送達された日時) を、`docdate` (E メールが送信された日付) の代わりに使用することもできます。例えば、`postedtime="2002-01-02 15:57"` は、2002 年 1 月 2 日午後 3:57 に送られた E メールを検索し、`deliveredtime>="2002-01-02 15:56" AND deliveredtime<"2002-01-02 15:58"` は、2002 年 1 月 2 日午後 3:56 から午後 3:58 の間に送達された E メールを検索します。

フィールド検索の日付は、ISO 8601 標準を使用して、YYYY-MM-DD または YYYY-MM-DD HH:MM のように指定します。

`>=`、`<=`、`=`、`>`、または `<` などの演算子を使用できます。`docdate`、`postedtime`、`deliveredtime` の後にコロン (:) を付ける必要はありません。これらは予約語なので、フィールド・ラベルとしてではなく、照会式として使用できます。

表 8. 日時検索の例

検索対象	検索照会ボックスに入力する照会の例
2007 年 12 月 18 日より前に送信された文書	<code>docdate &lt; "2007-12-18"</code>
2007 年 12 月 18 日より後に送信された文書	<code>docdate &gt; "2007-12-18"</code>
2006 年 5 月 16 日以降、2007 年 12 月 18 日までの間に送信された文書	<code>docdate &gt;= "2006-05-16" AND docdate &lt;= "2007-12-18"</code>
2002 年 1 月 2 日午後 3:56 から午後 3:58 の間に送達された E メール	<code>deliveredtime&gt;="2002-01-02 15:56" AND deliveredtime&lt;"2002-01-02 15:58"</code>

## 検索照会ボックスを使用する場合のヒント

特定のフィールド名および予約語を使用する検索照会ボックスで、検索照会を入力または編集できます。

### 検索フィールド

表 9. キーワードおよび件名の構文

検索フィールド	検索照会ボックスの構文	検索対象の E メールの部分	検索照会ボックスでの入力の例
キーワード	なし	件名、本文、添付ファイル	ストック・オプション
件名	subject:	件名のみ	subject:401k

- 検索項目が二重引用符で囲まれていない場合は、代替形式がサポートされていない場合を除き、代替形式がサポートされます。検索項目の順序は、重要ではありません。大/小文字は区別します。特殊文字および予約語を照会項目に使用する場合は、エスケープする必要があります。フィールドでラベル付けした照会の場合、空白文字で区切られた複数の検索項目は、各項目を括弧で囲んでグループ化する必要があります。
- 検索項目が二重引用符で囲まれている場合 ("phrase" 構文)、代替形式は使用不可です。検索項目の順序は重要であり、突き合わせでは、照会内の連続する項目が文書内の連続する位置に検出される必要があります。大/小文字は区別しません。少数の特殊文字のみが、照会項目として使用する場合にエスケープする必要があります (詳しくは照会構文のページを参照)。フィールドでラベル付けした照会の場合、空白文字で区切られた複数の検索項目は、二重引用符でグループ化されません。
- 「件名」フィールドではワイルドカードを使用できます。ワイルドカードでは、単語の区切り文字として使用される句読文字は突き合わせせず、単語の区切り文字を超えても突き合わせしません。ワイルドカードについては、詳細な構文規則を参照してください。

表 10. 送信者および受信者の構文

検索ボックス・フィールド	検索照会ボックスの構文	検索対象の E メールの部分	検索照会ボックスでの入力の例
送信者	from:	送信者、プリンシパル	from:"john smith" from:jsm?th@example.com from:jjones@example.com
受信者	to:cc:bcc:	to、cc、bcc のリスト	to:cc:bcc:"jill jones" to:cc:bcc:jjones@example.com

- ワイルドカードを使用できます。ワイルドカードでは、単語の区切り文字として使用される句読文字は突き合わせせず、単語の区切り文字を超えても突き合わせしません。ワイルドカードについては、詳細な構文規則を参照してください。
- 代替形式はサポートされません (原形に変換して検索されません)。
- このフィールドにアットマーク (@) が含まれるストリングはすべて、E メール・アドレスとみなされます。不完全な E メール・アドレスでは、索引付き Eメールの完全アドレスに一致しません。ワイルドカードを含むアドレスは、不完全とは限りません。例えば、John\*@example では John.Smith@example.com に一致し

ませんが、John\*@example.com では一致します。あるいは、Smith@example.com は example.com のすべての Smiths に一致しません。アドレスの両方の部分 (ID とドメイン) にワイルドカードを使用できます。例えば、John\*@example\* で John Smith@example.comを検出することができます。

- 検索フィールドの右側にあるアイコンをクリックした場合に表示されるダイアログに入力する検索項目に対しても、同じ注意事項が適用されます。

表 11. 日付フィールドの構文

検索フィールド	検索照会ボックスの構文	検索対象の E メールの部分	検索照会ボックスでの入力の例
日付の範囲	docdate	送信の日付	(docdate>="2008-08-12" AND docdate<="2008-08-13")

docdate の後にコロン (:) は必要ありません。 docdate は、照会式に使用される予約語であり、フィールド・ラベルではありません。代替形式はサポートされません。

## カテゴリー

表 12. カテゴリー

検索フィールド	検索照会ボックスの構文	検索対象の E メールの部分	検索照会ボックスでの入力の例
句	phrase:	件名、本文、添付ファイル	phrase:"case sensitive"
送信者	senders:	送信者、プリンシパル	senders:"john.smith@example.com"
受信者	recipients:	To, cc, bcc のリスト	recipients:"jill.jones@example.com"
送信者ドメイン	senderdomains:	送信者、プリンシパル	senderdomains:"example.com"
受信者ドメイン	recipientdomains:	To, cc, bcc のリスト	recipientdomains:"domain.example.com"
人物	person:	件名、本文、添付ファイル	person:"John Smith"
企業	company:	件名、本文、添付ファイル	company:IBM
ロケーション	location:	件名、本文、添付ファイル	location:"silicon valley"



検索項目の代替形式はサポートされません。

表 13. フラグ

検索フィールド	検索照会ボックスの構文	検索対象の E メールの部分	検索照会ボックスでの入力の例
フラグ	flags:	フラグ・テーブル	flags:"Confidential"

フラグ名は大/小文字が区別されます。代替形式はサポートされません。

## 検索照会ボックス

表 14. 送信者および受信者

検索フィールド	検索照会ボックスの構文	検索対象の E メールの部分	検索照会ボックスでの入力の例
なし	to:	to リストのみ	to:john.sm*th@example.com
なし	bcc:	bcc リストのみ	bcc:jill.jo*es@example.net
なし	cc:	cc リストのみ	cc:"john smith" OR cc:"jill jones"
なし	sentrepresenting:	プリンシパルのみ	sentrepresenting:"Jill Jones"

- 検索項目が二重引用符で囲まれていない場合は、注記 5 が適用される場合を除き、代替形式がサポートされます。検索項目の順序は、重要ではありません。大/小文字は区別しません。特殊文字および予約語を照会項目に使用する場合は、エスケープする必要があります。フィールドでラベル付けした照会の場合、空白文字で区切られた複数の検索項目は、各項目を括弧で囲んでグループ化する必要があります。
- 検索項目が二重引用符で囲まれている場合 ("phrase" 構文)、代替形式は使用不可です。検索項目の順序は重要であり、突き合わせでは、照会内の連続する項目が文書内の連続する位置に検出される必要があります。大/小文字は区別しません。少数の特殊文字のみが、照会項目として使用する場合にエスケープする必要があります (詳しくは照会構文のページを参照)。フィールドでラベル付けした照会の場合、空白文字で区切られた複数の検索項目は、二重引用符でグループ化されません。
- ワイルドカードを使用できます。ワイルドカードでは、単語の区切り文字として使用される句読文字は突き合わせせず、単語の区切り文字を超えても突き合わせしません。ワイルドカードについては、詳細な構文規則を参照してください。
- 代替形式はサポートされません。

表 15. 日時の予約語

検索フィールド	予約語	検索対象の E メールの部分	検索照会ボックスでの入力の例
なし	postedtime	E メールが送信された日時	postedtime="2002-01-02 15:57"
なし	deliveredtime	E メールが送達された日時	deliveredtime>="2002-01-02 15:56" AND deliveredtime<"2002-01-02 15:58"

postedtime または deliveredtime の後にコロン (:) は必要ありません。これらは照会式に使用される予約語であり、フィールド・ラベルではありません。

## 検索結果の解釈

検索結果を使用してケースの内容に関する理解を深め、その後結果を絞り込むことができます。

検索ページには以下が表示されます。

- 検索照会
- 検索結果の時刻表
- 検索結果のリスト
- 検索結果で検出されたカテゴリーの内容

### 検索照会ボックス

すべての照会は、編集可能な検索照会ボックスに表示されます。検索は保存して再利用が可能であり、各照会を前後に 1 つずつ移動することもでき、また照会を編集して新規の照会を実行することも可能です。

検索照会の構文規則を知っている場合は、テキスト・ボックスに照会を直接入力できます。

### 時刻表

時刻表には、検索結果の文書の数、文書が送信された日付、月、または年ごとにグラフィカルに表示されます。スライダーを移動して日付範囲を調整できます。このスライダーは、カーソルを時刻表に移動すると表示されます。「**結果の更新**」をクリックすると、検索結果内の日付範囲が新規の範囲に変更されます。

スライダーを使用して日付範囲を調整した場合は、新規に指定した日付範囲をクリックしてズームインできます。時刻表バー・グラフの垂直バーをクリックすると、その期間のみの結果が表示されます。

時刻表には、表示される各日付範囲ごとに、実際の文書数、および想定された文書数の両方が表示されます。想定された文書数は、対象の日付範囲において、ケース全体にわたって結果を均一に配分した場合を前提に計算されています。実際の数が想定された数より多い場合は、その期間でアクティビティが多いことを示しており、そのため検討者が注目することになります。例えば、句「セクシャル・ハラスメント」を検索し、ケース内の 500 万件の E メール の 5% でその句が一致した場



合、選択した任意の期間の E メール の 5% でその句が一致すると予想できます。このような状況において、2007 年 1 月に送信された E メール の 20% でその句が検出されたとしたら、セクシャル・ハラスメントの問題に関してその期間の E メールを検討することになると考えられます。

## カテゴリー

検索エンジンでは、文書内容を評価し、テキストの分析規則に従って語および句を分類します。検索結果の語および句は、カテゴリーに表示されます。各カテゴリーの語および句は、頻度別にソートされます。検索照会では、選択した語または句を追加または除外できます。

「詳細表示」をクリックした場合に表示される各カテゴリーの拡張リストには、現在の検索結果での E メール数に対する各カテゴリー内の語および句の相関が小さい棒グラフで表示されます。数が多い場合は、その語または句が結果に多く出現していることを表しています。数が少ない場合は、結果でのその語または句の出現が少ないことを表しています。

### 関連概念

1 ページの『検索ストラテジー』

### 関連タスク

『カテゴリーによる検索結果の絞り込み』

---

## カテゴリーによる検索結果の絞り込み

検索結果は、検索カテゴリーを使用して絞り込むことができます。

カテゴリーを使用して検索結果を絞り込むには、以下のようになります。

1. 「検索」ボックスで、1 つ以上のカテゴリーを展開します。
2. 以下のように照会を変更します。

目的	実行する操作
選択した検索項目を現在の検索照会に即座に追加する	検索項目をクリックします。  ブールの AND 演算子を使用して、選択した検索項目を組み込むことにより、結果を絞り込む新規の検索が実行されます。
現在の検索照会に 1 つ以上の検索項目を組み込む	選択した検索項目の緑色のプラス符号 (+) をクリックし、「検索」をクリックします。  カテゴリーを使用した検索時に、「検索への追加」が自動的に選択されます。選択した検索項目のみで新規の検索を実行するには、「新規検索」を選択し、「検索」をクリックします。

目的	実行する操作
現在の検索照会から 1 つ以上の検索項目を除外する	<p>選択した検索項目の赤色のマイナス符号 (-) をクリックし、「検索」をクリックします。</p> <p>カテゴリーを使用した検索時に、「検索への追加」が自動的に選択されます。選択した検索項目のみで新規の検索を実行するには、「新規検索」を選択し、「検索」をクリックします。</p>

- 「句」などの同じカテゴリー内、および同じ符号で選択された検索項目は、ブールの OR 演算子を使用して検索照会に組み込まれます。
- 赤色のマイナス符号 (-) を使用して選択された同じカテゴリー内の検索項目は、ブールの AND NOT 演算子を使用して検索照会に組み込まれます。
- 「句」および「送信者」などの異なるカテゴリー内で選択された検索項目は、ブールの AND 演算子を使用して検索照会に組み込まれます。

例えば、同じカテゴリー内で複数の緑色のプラス符号 (+) を選択した場合は、`phrase:("legal" OR "illegal")` などの照会が生成されます。同じカテゴリー内で複数の赤色のマイナス符号 (-) を選択した場合は、`phrase:(NOT ("acceptable" OR "legal"))` などの照会が生成されます。

照会 `country` では、「句」カテゴリーにメキシコ、イタリア、タンザニア、およびインドなどの結果が戻され、「企業」カテゴリーには IBM、Tata、および Toyota などの結果が戻されます。インドの横にある緑色のプラス符号 (+)、イタリアの横にある赤色のマイナス符号 (-)、および Tata の横にある緑色のプラス符号 (+) をクリックし、「検索」をクリックすると、結果の検索照会は `(country AND (phrase: (India AND NOT Italy))) AND (company: Tata)` になります。

#### 関連概念

16 ページの『検索結果の解釈』

## 検索の保存

実行した検索照会の保存、保存した検索の編集、保存した検索のオープンと実行、および検索のインポートが可能です。

検索照会を保存するには、その検索を検索照会ボックスに表示する必要があります。検索照会ボックスに照会が表示されていない場合は、検索フィールドに検索項目を入力して検索を実行してください。

保存した検索は、ユーザーおよびケースに関連付けられます。システム全体で保存される検索はありません。保存された検索を別のユーザーと共有するには、検索照会をコピーして、テキストとして別のユーザーに送信します。

ケースのページを開くたびに実行されるデフォルト検索として、保存された検索を選択できます。

保存された検索はグループ化して、関連する検索をまとめておくことができます。

検索照会を保存するには、以下のようになります。

1. 「検索」 ページで、検索照会を入力または実行します。
2. 「保存」 をクリックします。
3. 検索照会を保存する先の、保存済み検索のグループを指定または選択します。
4. 「保存」 をクリックします。

## 保存された検索の管理

保存された検索は、実行、編集、または削除が可能です。また、保存された検索をデフォルトの検索照会として指定することもできます。

保管済み検索を管理するには、以下のようにします。

1. 検索ページで、「**保存された検索**」 をクリックします。
2. 以下の操作を選択します。

目的	実行する操作
保存された検索の実行	保存された検索を選択して「 <b>実行</b> 」 をクリックするか、保存された検索をダブルクリックします。
保存された検索の編集	保存された検索を選択し、グループまたは照会を編集して、「 <b>保存</b> 」 をクリックします。
検索の削除	削除する対象の保存された検索に対して「 <b>削除</b> 」 をクリックします。

3. オプション: 選択した検索照会をデフォルトとして指定するには、「**デフォルトの検索**」 を選択します。デフォルト検索は、ケースの検索ページを開いたとき、または検索インターフェースで「**リセット**」 をクリックしたときに自動的に実行されます。

## 保存された検索のエクスポート

保存された検索はエクスポートして、別の検討者と共有したり、別のケースで使用したりできます。

検索照会はテキスト・ファイルからのインポートが可能であり、保存された検索をテキスト・ファイルにエクスポートできます。保存された検索のテキスト・ファイルは、1 行ごとに 1 つの照会のフォーマットにする必要があります。

保存された検索をエクスポートするには、以下のようにします。

1. 「検索」 ページで、「**保存された検索**」 をクリックします。
2. エクスポートする保存済み検索を選択します。
3. 「**エクスポート**」 をクリックします。
4. 保存された検索のファイル名およびロケーションを指定します。

## 検索のインポート

別の検討者またはケースからエクスポートした保存済み検索をインポートできます。

検索照会はテキスト・ファイルからのインポートが可能であり、保存された検索をテキスト・ファイルにエクスポートできます。テキスト・ファイルは、サポートされている構文を使用した検索照会を 1 行ごとに 1 つ入力するフォーマットにします。

検索をインポートするには、以下のようにします。

1. 「検索」 ページで、「**保存された検索**」をクリックします。
2. 「**保存された検索のインポート**」をクリックします。
3. 保存された検索が含まれているテキスト・ファイルを指定または参照します。
4. 検索照会を保存する先の、グループを指定または選択します。 グループを指定しない場合、検索は「**グループ化されていない**」グループに保存されます。
5. 「**インポート**」をクリックします。

---

## 第 3 章 フラグによる文書の関連性の識別

各ケースのディスカバリー要求の各要件への関連性ごとに、選択した文書を識別するにはフラグを設定します。ケース管理者がアーカイブにケースを戻すために転送すると、各文書はフラグによってグループ化できます。

フラグはメタデータであり、1 つ以上の文書に関連付けられますが、文書の一部ではありません。

例えば、E メール文書がディスカバリー要求に反応すると考えられる場合は、それを「反応あり」としてフラグ付けします。ケース内のすべての E メール文書を検討し、文書のサブセットへのフラグ付けが完了したら、「反応あり」のフラグがあるすべての文書を検索して、法律顧問による検討のためにグループ化できます。または、明らかに無関係である E メール文書にフラグを立て、後続の検索からそれらの文書を除外できます。

文書に設定したフラグにより、追加検討用に送られる文書が最終的に特定されます。フラグ付けされた文書は、管理者が決定するアーカイブ内の新規フォルダーに保管できます。

一般的なフラグ・セットはデフォルトとして組み込まれています。管理者は、ユーザーのビジネス要件に合致するように、デフォルトの一連のフラグに対してフラグを追加または削除したり、ケースにフラグを追加したりできます。また、管理者は既存のフラグを編集することもできます。フラグはケース管理者によってカスタマイズ可能であるため、管理者と連携して、特定のケースで使用するフラグを決定する必要があります。

デフォルトのフラグは以下のとおりです。

- 反応あり
- 反応なし
- 特権
- 最重要
- 作成なし
- 機密情報
- 第 2 レベルの検討
- 外国語

初期検索後に、結果リスト内の文書に設定されたフラグが「フラグ」カテゴリーに表示され、選択してさらに検索結果を絞ることができます。また、特定のフラグを選択して、実行する検索のすべてに含める、または除外することもできます。

フラグを設定するには、以下のようにします。

1. 「結果リスト」で、1 つ以上の文書を選択します。1 つの文書、または複数の文書を選択するか、「すべて選択」チェック・ボックスを選択します。
2. 「フラグ」をクリックします。

3. フラグを 1 つ以上選択し、「**保存**」をクリックします。 選択された E メールに既にフラグが設定されている場合は、設定を変更できます。また、文書の「**n フラグ**」のリンクをクリックすることにより、個々の文書のフラグを表示または設定することもできます。

#### **関連タスク**

5 ページの『第 2 章 ケースの検索』

---

## 特記事項

本書は米国 IBM が提供する製品およびサービスについて作成したものです。

本書に記載の製品、サービス、または機能が日本においては提供されていない場合があります。日本で利用可能な製品、サービス、および機能については、日本 IBM の営業担当員にお尋ねください。本書で IBM 製品、プログラム、またはサービスに言及していても、その IBM 製品、プログラム、またはサービスのみが使用可能であることを意味するものではありません。これらに代えて、IBM の知的所有権を侵害することのない、機能的に同等の製品、プログラム、またはサービスを使用することができます。ただし、IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

IBM は、本書に記載されている内容に関して特許権 (特許出願中のものを含む) を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

〒106-8711  
東京都港区六本木 3-2-12  
日本アイ・ビー・エム株式会社  
法務・知的財産  
知的財産権ライセンス渉外

**以下の保証は、国または地域の法律に沿わない場合は、適用されません。** IBM およびその直接または間接の子会社は、本書を特定物として現存するままの状態を提供し、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。

この情報には、技術的に不適切な記述や誤植を含む場合があります。本書は定期的に見直され、必要な変更は本書の次版に組み込まれます。IBM は予告なしに、随時、この文書に記載されている製品またはプログラムに対して、改良または変更を行うことがあります。

本書において IBM 以外の Web サイトに言及している場合がありますが、便宜のため記載しただけであり、決してそれらの Web サイトを推奨するものではありません。それらの Web サイトにある資料は、この IBM 製品の資料の一部ではありません。それらの Web サイトは、お客様の責任でご使用ください。

IBM は、お客様が提供するいかなる情報も、お客様に対してなんら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。



本プログラムのライセンス保持者で、(i) 独自に作成したプログラムとその他のプログラム (本プログラムを含む) との間での情報交換、および (ii) 交換された情報の相互利用を可能にすることを目的として、本プログラムに関する情報を必要とする方は、下記に連絡してください。

IBM Corporation  
J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003  
U.S.A.

本プログラムに関する上記の情報は、適切な使用条件の下で使用することができませんが、有償の場合もあります。

本書で説明されているライセンス・プログラムまたはその他のライセンス資料は、IBM 所定のプログラム契約の契約条項、IBM プログラムのご使用条件、またはそれと同等の条項に基づいて、IBM より提供されます。

この文書に含まれるいかなるパフォーマンス・データも、管理環境下で決定されたものです。そのため、他の操作環境で得られた結果は、異なる可能性があります。一部の測定が、開発レベルのシステムで行われた可能性があります。その測定値が、一般に利用可能なシステムのものと同じである保証はありません。さらに、一部の測定値が、推定値である可能性があります。実際の結果は、異なる可能性があります。お客様は、お客様の特定の環境に適したデータを確かめる必要があります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者をお願いします。

IBM の将来の方向または意向に関する記述については、予告なしに変更または撤回される場合があります、単に目標を示しているものです。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名称はすべて架空のものであり、名称や住所が類似する企業が実在しているとしても、それは偶然にすぎません。

著作権使用許諾:

本書には、様々なオペレーティング・プラットフォームでのプログラミング手法を例示するサンプル・アプリケーション・プログラムがソース言語で掲載されています。お客様は、サンプル・プログラムが書かれているオペレーティング・プラットフォームのアプリケーション・プログラミング・インターフェースに準拠したアプリケーション・プログラムの開発、使用、販売、配布を目的として、いかなる形式においても、IBM に対価を支払うことなくこれを複製し、改変し、配布することができます。このサンプル・プログラムは、あらゆる条件下における完全なテストを



経ていません。従って IBM は、これらのサンプル・プログラムについて信頼性、利便性もしくは機能性があることをほのめかしたり、保証することはできません。

それぞれの複製物、サンプル・プログラムのいかなる部分、またはすべての派生した創作物にも、次のように、著作権表示を入れていただく必要があります。「© (お客様の会社名) (西暦年). このコードの一部は、IBM Corp. のサンプル・プログラムから取られています。© Copyright IBM Corp. 2004, 2006. All rights reserved.

この製品の一部は、著作権により保護されています。

- Copyright © Outside In® Viewer Technology 1992-2005 Stellent Chicago, Inc. All rights reserved.

## 商標

以下は、International Business Machines Corporation の米国およびその他の国における商標です。 <http://www.ibm.com/legal/copytrade.shtml>

Microsoft、Windows および Windows NT は、Microsoft Corporation の米国およびその他の国における商標です。

Java およびすべての Java 関連の商標およびロゴは Sun Microsystems, Inc. の米国およびその他の国における商標です。

Adobe、Adobe ロゴ、PostScript、PostScript ロゴは、Adobe Systems Incorporated の米国およびその他の国における登録商標または商標です。

他の会社名、製品名およびサービス名等はそれぞれ各社の商標です。



## 用語集

この用語集を使用して、検索システムのコセを理解するのに役立ててください。

### アクセス制御リスト (ACL) (access control list (ACL))

コンピューター・セキュリティーにおける、オブジェクトに関連付けられたリスト。このリストにより、オブジェクトにアクセス可能なすべての対象とそのアクセス権を識別する。

### アノテーター (annotator)

UIMA アノテーターは、非構造化情報を UIMA フレームワークの一部として解析する。「UIMA」も参照。

### 開始 Web アドレス (starting Web address)

Web クロールの開始点。開始 Web アドレス (www.example.org など) を入力すると、クローラーは、開始ページから HTML リンクをたどることによって到達可能な、そのサイトにあるすべての Web ページ (文書) をクロールする。

### 開始ディレクトリー (starting directory)

ファイル・システム・クロールの開始点。開始ディレクトリー (C:\mydocuments など) を入力すると、クローラーは、サブディレクトリーを含めて、そのディレクトリー内のすべての文書をクロールする。

### 鍵ストア・ファイル (keystore file)

署名者の証明書として保管された公開鍵と、個人証明書に保管された秘密鍵の両方が入っている鍵リング。

### 鍵リング (key ring)

コンピューター・セキュリティーにおいて、公開鍵、秘密鍵、トラステッド・ルート、および証明書が入っているファイル。「鍵ストア・ファイル (keystore file)」も参照。

### キーワード一致ランキング (keyword match ranking)

結果ごとに、検索エンジンは、結果がどの程度正確に照会と一致しているかを判別する。文書に含まれた照会語の出現数が多い

ほど、また、出現する複数のキーワードが互いに類似しているほど、結果が結果リスト内に表示される可能性が高くなる。キーワードの一致は、検索結果を返すための最も重要な要因である。

### クローラー (crawler)

データ・ソースまたはコンテンツ・リポジトリーから文書を取得し、検索索引の作成に使用できる情報を収集するソフトウェア・プログラム。

### クロール (crawling)

コンテンツ・リポジトリーやファイル・システムなどのデータ・ソースから文書を取得するソフトウェア・プログラムのアクティビティー。取得された文書はこの後、検索エンジンによって処理され、解析され、索引付けされて、ユーザーによる検索を可能にする。「クローラー (crawler)」および「クロール・スペース (crawl space)」も参照。

### クロール・スペース (crawl space)

索引付けの項目を取得するためにクローラーが読み取る、指定されたパターン (例えば、URL、データベース名、コンテンツ・リポジトリー、ファイル・システム・パス、ドメイン・ネーム、IP アドレスなど) に一致するソースのセット。

### 言語学的検索 (linguistic search)

基本型に変換された語 (例えば、*mice* が *mouse* として索引付けされるようにしたり、複合語を拡張したりする) で文書をブラウズ、取得、および索引付けする検索タイプ。

### 検索エンジン (search engine)

検索要求を受け入れ、文書のリストをユーザーに返すプログラム。

### 検索キャッシュ (search cache)

以前の検索要求のデータと結果を保存するバッファー。

## 検索結果 (search results)

検索要求に一致する文書のリスト。

## 検索ページ (search page)

ユーザー照会を受け入れ、検索結果のリストを表示する Web ページ。

## 索引 (index)

照会語を含む文書を見つける検索を使用可能にするためのデータ項目を参照するデータ構造。

## 資格情報 (credential)

ユーザー、グループの関連付け、およびその他のセキュリティー関連の識別属性を記述する、認証中に取得される詳細な情報。資格情報は、許可、監査、代行など、多数のサービスを実行するために使用できる。例えば、ユーザーのサインオン情報 (ユーザー ID とパスワード) は、そのユーザーにアカウントへのアクセスを許可する資格情報である。

## 字句類縁性 (lexical affinity)

文書において意味上、互いに近い検索語の関係。字句類縁性は、結果の適合性を計算するのに使用される。

## システム・ログ (system log)

検索エンジンの実行と並行したコンピューター・アクティビティーの記録。システム・ログを調べると、エンタープライズ検索システム上の問題を確認できる。

## シノニム (synonym)

同じ意味を持った異なる語。照会の中にシノニムの 1 つが出現した場合に、検索エンジンが検索対象を広げるために使用するシノニムを定義することによって、ユーザーが正しい文書を検出できる可能性が高まる。

## 照会ログ (query log)

照会アクティビティーの記録。照会ログを調べると、どのユーザーが検索しているか、照会にどの程度の時間がかかっているか、およびその他の情報を見つけることができる。

## 証明書 (certificate)

コンピューター・セキュリティーにおいて、公開鍵を証明書所有者の ID にバインドし、証明書所有者の認証を可能にするデ

ジタル文書。証明書は、認証局によって発行され、その認証局によってデジタル署名される。

## ステミング (stemming)

「ワード・ステミング (word stemming)」を参照。

## ストップワード (stop word)

検索アプリケーションによって無視される *the*、*an*、*and* など、広く使用される語。

## ストップワードの除去 (stop word removal)

よくある語を無視し、より適切な結果を返すために、照会からストップワードを除去するプロセス。

## 正規化 (normalization)

「文字の正規化 (normalization)」を参照。

## セグメンテーション (segmentation)

別個の字句単位へのテキストの分割。非辞書ベースの処理には、空白文字および n-gram セグメンテーションがあるのに対して、辞書ベースのサポートには、語、文、および段落セグメンテーション、および見出し語を決めるための分類整理がある。

## 接語 (clitic)

構文的には独立して機能するが、発音上は他の語に結びつく語。接語は、結びつけられた語に連結、または分離して書くことができる。接語の一般的な例としては、英語における短縮形の最後の部分 (*wouldn't* または *you're*) などがある。

## 説明付きリンク (featured links)

ユーザーが特定の照会を入力するたびに、検索ページの上部に表示されるように構成されるタイトルと説明を伴うリンク。説明付きリンクには、照会、Web アドレス、タイトル、およびオプション・サマリーの 4 つの部分がある。

## ソフト・エラー・ページ (soft error page)

要求された Web ページがなぜ返されないかについての情報を提供する Web ページのタイプ。例えば、単純な状況コードを返すのではなく、HTTP サーバーは、状況コードを詳細に説明するページを返すことができる。

### データ・ソース (data source)

Web、ファイル・システム、データベースなど、文書の取得元となる、データのリポジトリ。

### ディレクトリーの深さ (directory depth)

ファイル・システム・パスにおけるサブディレクトリーの長さまたは数。例えば、ディレクトリー C:\My Company\personnel\private\records は、C:\My Company\personnel よりも深い (より多くのサブディレクトリーがある)。「URL の深さ (URL depth)」も参照。

### テキスト・セグメンテーション (text segmentation)

「セグメンテーション (segmentation)」を参照。

### トークナイザー (tokenizer)

テキストをスキャンし、一連の文字をトークンとして認識するかどうか、およびそれをいつ行うかを判別するテキスト・セグメンテーション・プログラム。

### トークン (token)

エンタープライズ検索で索引付けされる基本テキスト単位。トークンは、各言語における語の場合もあれば、索引付けに適切なテキストから構成される他の単位となる場合もある。

### トークン化 (tokenization)

入力データをトークンに構文解析するプロセス。

### 認証 (authentication)

ユーザーまたはサーバーの ID を検証するプロセス。

### 認証局 (certificate authority)

デジタル署名、および公開鍵と秘密鍵のペアの作成に使用されるデジタル証明書を発行する、信頼できるサード・パーティーの組織または企業。認証局は、固有の証明書を付与された個人の ID を保証する。

### ノーマライザー (normalizer)

テキストをスキャンし、大文字化および発音区別符号など、文字の変化形を共通形式に変換する、文字の正規化プログラム。

### パーサー (parser)

エンタープライズの検索データ・ストアに追加された文書を解釈するプログラム。パーサーは、文書から情報を抽出し、その情報を索引作成、検索、および取得のために準備する。

### 発音区別符号 (diacritic)

1 つの文字、または文字の組み合わせの音価の変化を示すマーク。

### パラメトリック検索 (parametric search)

日付、整数、その他の、指定された範囲内の数値データ・タイプなど、数字による値または属性を含むオブジェクトを探す検索のタイプ。

### 非索引付けディレクティブ (no-index directive)

その Web ページのコンテンツを索引に含めないよう、ロボット (Web クローラーなど) に命令する、Web ページ内のディレクティブ。

### 非追跡ディレクティブ (no-follow directive)

その Web ページ内にあるリンクをたどらないよう、ロボット (Web クローラーなど) に命令する、Web ページ内のディレクティブ。

### ファイル・システム・クローラー (File system crawler)

ディレクトリーから文書を取得するクローラーのタイプ。

### フィールド (field)

データの特定のカテゴリー、または制御情報を入力する領域。

### フィールド限定の検索 (fielded search)

特定のフィールドに制限された照会。

### フリー・テキスト検索 (free text search)

検索語がフリー・フォーム・テキストとして表現される検索。

### フリー・フォーム・テキスト (free-form text)

語または文で構成された、非構造化テキスト。

### 変更日付ランキング (modification date ranking)

多くの場合、新しい文書は、古い文書よりも高くランクされる。

変更日付によるランキングは、多くのタイプの文書にとって重要となる。ただし、多

くの Web サーバーによって返される最終変更日は、文書の正確な経過日数を反映しないため、文書日付は、Web サイト文書の場合、通常は信頼できない。

#### マスク文字 (masking character)

検索語の先頭、中央、および末尾で任意の文字を表すために使用される文字。マスク文字は通常、索引で語の変化形を見つけるために使用される。「ワイルドカード文字 (wildcard character)」も参照。

#### 見出し語 (lemma)

語の基本型。見出し語は、チェコ語など、強く語形変化した言語において重要である。

#### 見出し語を決めるための分類整理 (lemmatization)

語の原形および異なる文法的な形を識別するプロセス。例えば、mouse の検索では、語 mice を含む文書も検出され、go の検索では、going、gone、または went を含む文書も検出される。

#### 無視するテキスト (ignore text)

検索照会で無視されるユーザー定義のテキスト (定形文面の特記事項など)。ストップワードと類似。「ストップワード (stop word)」も参照。

#### メタデータ (metadata)

特定の情報を記述し、その情報を検索において取得し、カテゴリーにおいてブラウズし、または目的に応じてフィルターするのに役立つデータ。メタデータは、多くの場合、分類法または分類スキームの一部である。

#### 文字の正規化 (character normalization)

大文字化および発音区別符号などの、文字の変化形を共通形式に変換するプロセス。

#### ユーザー・エージェント (user agent)

Web をブラウズし、訪れたサイトに自身の情報を残すアプリケーション。エンタープライズ検索においては、Web クローラーはユーザー・エージェントである。

#### ユーザー・エージェント・ストリング (User-Agent string)

Web クローラーを識別する。使用法の 1 つとして、特定のユーザー・エージェント・ストリングに基づいてアクセスを拒否できる robots.txt ファイルが該当する。

#### ランキング (ranking)

照会からの検索結果における各文書に対する整数値の割り当て。検索結果における文書の順序は、照会に対する適切さに基づく。ランクが高いほど、一致が近いことを意味する。

#### ワード・ステミング (word stemming)

語の変化形が共通形式に変換される、言語学的正規化のプロセス。例えば、connections、connective、connected などの語は connect に変換される。

#### ワイルドカード文字 (wildcard character)

検索語の先頭、中央、または末尾で任意の文字を表すために使用される文字。

#### GET コマンド (GET command)

Web サーバーからファイルを要求する HTTP コマンド。

#### HTML フォーム・ベースの認証 (HTML form-based authentication)

ユーザー資格情報を提供するための HTML ページでフォームを直接使用する認証方式。

#### HTML メタ・タグ (HTML meta tags)

HTML メタ・タグは、HTML 文書に追加して、例えば、クローラーに指示を与えることができる。

#### HTTP 基本認証 (HTTP basic authentication)

HTTP プロトコルで指定され、Web サイトへのアクセスを制御するために設計された標準認証スキーム。HTTP 基本認証では、ユーザー名とパスワードのみが使用される。

#### HTTP プロキシ・サーバー (HTTP proxy server)

アプリケーションまたは Web サーバーによってホストされる HTTP Web 要求の仲介プログラムとして機能するサーバー。プロキシ・サーバーは、エンタープライズ内のコンテンツ・サーバーの代理の役割を果たす。

#### IP アドレス (IP address)

IP 標準が使用されるネットワーク上のデバイスまたは論理装置の固有のアドレス。

#### Java Database Connectivity (JDBC)

Java プラットフォームと、広範囲にわたるデータベースとの間の、データベースに



依存しない接続のための業界標準。JDBC インターフェースは、SQL ベースのデータベース・アクセスのために、呼び出しレベルの API を提供する。

#### **Lightweight Directory Access Protocol (LDAP)**

TCP/IP を使用して、X.500 モデルをサポートするディレクトリーへのアクセスを提供し、より複雑な X.500 Directory Access Protocol (DAP) のリソース要件を適用する必要がない公開プロトコル。例えば、LDAP を使用すると、インターネットまたはイントラネットのディレクトリー内の人物、組織、およびその他のリソースを見つけることができる。

#### **MIME タイプ (MIME type)**

インターネットを介して転送されるオブジェクトのタイプを識別するインターネット標準。

#### **n-gram セグメンテーション (n-gram segmentation)**

語を区切るのにブランク・スペースまたは句読法を使用する方式 (Unicode ベースの空白文字セグメンテーションなど) ではなく、特定の数の文字のオーバーラップ・シーケンスを単一の語と見なす分析方式。例えば、 $n=2$  である場合、テキスト ABEFD は、シーケンス AB BE EF FD に区切られる。

#### **POST コマンド (POST command)**

情報を処理のために Web サーバーに送信する HTTP コマンド。POST メソッドは、入力されたデータが入っているフォームをサーバーに送信するため、HTML ファイルにおいて広く実装されている。

#### **Robots Exclusion Protocol**

Web サイト管理者が、サイトを訪れるロボット (Web クローラーなど) にサイトの一部へのアクセスの禁止を指示できるプロトコル。

#### **robots.txt ファイル (robots.txt file)**

クローラーによるサーバーへのクローリングを阻止するため、一部の Web サイト管理者は、クローラー用のアクセス・ポリシーを定義するファイルを Web サーバー上に作

成する。このファイルは、robots.txt と呼ばれ、Robots Exclusion Protocol に準拠する。

#### **Secure Sockets Layer (SSL)**

通信プライバシーを提供するセキュリティ・プロトコル。SSL を使用すると、クライアント/サーバーのアプリケーションは、盗聴、改ざん、およびメッセージ偽造を防ぐために設計された方法で通信することができる。

**UIMA** Unstructured Information Management Architecture は、オープン・ソース・フレームワークであり、大量の非構造化情報において適切な情報を検出するアプリケーションを開発するための SDK である。「Apache UIMA」も参照。

#### **Unicode ベースの空白文字セグメンテーション (Unicode-based white space segmentation)**

Unicode 文字プロパティを使用して、トークン文字と区切り文字とを区別するトークン化の方式。

#### **Uniform Resource Identifier (URI)**

抽象リソースと物理リソースを識別する、短縮ストリングの文字。

#### **Uniform Resource Locator (URL)**

インターネットなどのネットワーク内でアクセス可能な情報リソースの固有のアドレス。URL には、情報リソースにアクセスするために使用されるプロトコルの省略名と、情報リソースを見つけるためにプロトコルによって使用される情報が含まれている。

#### **URL の深さ (URL depth)**

Web サイト・アドレスにおけるスラッシュの繰り返しの長さまたは数。例えば、Web サイト・アドレス `www.example.org/personnel/private/records` は、Web サイト・アドレス `www.example.org/personnel` よりも深い (より多くのスラッシュがある)。

通常、文書の URL が深くなるにつれて、一般的な関心度は低くなる。「ディレクトリーの深さ (directory depth)」も参照。

**Web クローラー (Web crawler)**

Web 文書を取得し、その文書内のリンクをたどることによって Web を探索するクローラーのタイプ。

**Web リンク分析ランキング (Web links analysis ranking)**

多くのリンクによってポイントされている文書が、リンクの少ない文書よりも高くランクされるランキングの方式。



## 索引

日本語, 数字, 英字, 特殊文字の順に配列されています。なお, 濁音と半濁音は清音と同等に扱われています。

### [ア行]

アルファベット以外の文字 9  
エスケープ・コード 9

### [カ行]

外国語フラグ 21  
概要 1  
カテゴリ 5, 10, 16  
    企業 17  
    句 17  
    受信者 17  
    受信者ドメイン 17  
    送信者 17  
    送信者ドメイン 17  
    人物 17  
    フラグ 17  
    ロケーション 17  
可変 8  
完全一致突き合わせ 6  
企業 10  
企業カテゴリ 14  
機密情報フラグ 21  
句 10  
句カテゴリ 14  
句検索 6  
ケース 5  
結果 16  
検索 16  
    カテゴリ 17  
    ストラテジー 1  
    フラグ 21  
    保存された検索 18  
検索演算子  
    接近性 6  
    ブール演算子 6  
    ワイルドカード 6  
検索照会の構文  
    句 6  
    ブール演算子 6  
    フィールド 6  
    用語の除外 6  
    用語の包含 6  
    ワイルドカード文字 6

検索照会ボックス 5  
検索のヒント 6, 13  
検索フィールド 5  
件名 10  
構文 13

### [サ行]

最重要フラグ 21  
作成なしフラグ 21  
時間 12  
時刻表 16  
受信者 10, 11  
受信者カテゴリ 14  
受信者ドメイン 10  
受信者ドメイン・カテゴリ 14  
照会 13, 16  
接近性 7  
送信者 10, 11  
送信者カテゴリ 14  
送信者ドメイン 10  
送信者ドメイン・カテゴリ 14

### [タ行]

第 2 レベルの検討フラグ 21  
特殊文字 9  
特権フラグ 21

### [ハ行]

反応ありのフラグ 21  
反応なしのフラグ 21  
日付 12  
人物 10  
人物カテゴリ 14  
ブール演算子 8  
ファセット・ブラウザ 17  
フィールド 10  
    キーワード 13  
    企業 13  
    句 13  
    件名 13  
    受信者 13  
    受信者ドメイン 13  
    送信者 13  
    送信者ドメイン 13  
日付 13  
人物 13  
フラグ 13

フィールド (続き)  
    ロケーション 13  
フラグ 10  
    外国語 21  
    機密情報 21  
    最重要 21  
    作成なし 21  
    設定 21  
    第 2 レベルの検討 21  
    特権 21  
    反応あり 21  
    反応なし 21  
    リスト 21  
フラグの設定 21  
フラグ・カテゴリ 14  
保存された検索 18  
    インポート 20  
    エクスポート 19  
    管理 19

### [マ行]

文字, 特殊 9

### [ラ行]

ロケーション 10  
ロケーション・カテゴリ 14

### [ワ行]

ワイルドカード 8  
ワイルドカード文字 6

## A

AND 8

## B

bcc: 11

## C

cc: 11

## **D**

deliveredtime 12  
docdate 12

## **F**

from: 11

## **I**

INORDER 7

## **N**

NOT 6, 8

## **O**

OR 8

## **P**

postedtime 12

## **T**

to: 11  
to: 11

## **W**

WITHIN 7





プログラム番号: 5724-V36

SC88-5795-00



日本アイ・ビー・エム株式会社  
〒106-8711 東京都港区六本木3-2-12