

## NTCIR-4における類似特許検索テストコレクションの構築

藤井 敦<sup>†</sup> 岩山 真<sup>††</sup> 神門 典子<sup>†††</sup>

<sup>†</sup> 筑波大学 図書館情報学系

<sup>††</sup> 東京工業大学 / 日立製作所

<sup>†††</sup> 国立情報学研究所

E-mail: fujii@slis.tsukuba.ac.jp

情報検索用テストコレクションを構築する動きが国際的に活発化し、新聞記事、論文抄録、Web文書などを対象にした大規模なコレクションが構築され、研究利用のために公開されている。特許検索システムは長い歴史を持つアプリケーションであるにもかかわらず、特許文書は従来の研究対象として扱われることが少なかった。本稿は、国立情報学研究所が主催する NTCIR-4 ワークショップにおいて遂行中の特許検索タスクと構築中のテストコレクションについて説明する。NTCIR-4 ワークショップは大学や企業などからシステム参加を募り、情報検索や自然言語処理に関する技術の評価を行っている。特許検索タスクでは、特許庁審査官やサーチャーが行う無効資料調査を代行する検索システムの評価を行っている。特許請求項を入力して、そこで請求されている権利を無効化できる別の特許を検索することが目的である。さらに、特許マップを自動作成するシステムの評価も行っている。ある技術に関して検索された特許を開発課題や達成手段など複数の軸で分類し可視化することが目的である。最終的に構築されるテストコレクションは、特許検索だけでなく、一般の情報検索に関する様々な研究課題を探求するために有効である。

### Building a Test Collection for Associative Patent Retrieval in NTCIR-4

Atsushi Fujii<sup>†</sup>, Makoto Iwayama<sup>††</sup>, Noriko Kando<sup>†††</sup>

<sup>†</sup>Institute of Library and Information Science, University of Tsukuba

<sup>††</sup> Tokyo Institute of Technology / Hitachi Ltd.

<sup>†††</sup> National Institute of Informatics

E-mail: fujii@slis.tsukuba.ac.jp

Large test collections targeting newspaper articles, technical abstracts, and Web documents have been produced by various international activities and are available for information retrieval research purposes. While a number of commercial patent retrieval systems and services have operated for a long time, patent documents have not received much attention in the research communities. In this paper we describe the Patent Retrieval Task, which is a part of the NTCIR-4 Workshop organized by the National Institute of Informatics, and the test collections produced in this task. The NTCIR-4 Workshop provides a forum in which researchers of universities and private companies share the same evaluation framework for information retrieval and natural language processing. We perform the invalidity search task, in which each participant group searches a patent collection for the patents that can invalidate the demand in an existing claim. We also perform automatic patent map generation, in which the patents associated with a specific topic are organized in a multi-dimensional matrix. The resultant test collections can be useful not only for patent retrieval purposes, but also for exploring various retrieval problems.

## 1 はじめに

再利用可能な情報検索用テストコレクションを構築する動きが1990年代から国際的に活発化し、TREC<sup>1</sup>、CLEF<sup>2</sup>、NTCIR<sup>3</sup>などのシステム参加型ワークショップで様々なコレクションが構築されてきた。その結果、実システムに要求される規模と比べても遜色ないほど大量の文書データが、研究利用のために公開されている。従来のテストコレクションで検索対象となるは新聞記事、論文抄録、Web文書が中心である。

特許検索システムは長い歴史を持つ商用サービス(アプリケーション)であるにもかかわらず、特許文書は従来の研究対象として扱われることが少なかった。しかし、特許文書(明細書や公報)は、以下に示すような点において、情報検索や自然言語処理などの学術的な研究対象としても興味深い。

- 発明に関する新語や専門用語を多く含む。
- 単件あたりの文書長が長い(ばらつきがある)。
- 権利を主張する部分である「請求項」が一種の制限言語で記述されている。
- テキストコンテンツ以外にも多数のフィールド(分類番号など)で構造化されている。
- 単一の発明内容について、抽象的な表現で書かれた請求項と具体的に書かれた詳細説明が同一文書内に存在する。
- 言語が異なる複数国への同時出願が行われるたびに、対訳コーパスが生成される。

研究対象として特許を扱い関連技術を発展させ、それらを社会に還元することは価値があると考えられる。

このような背景を踏まえて、筆者らは、国立情報学研究所が主催する第3回 NTCIR ワークショップ(NTCIR-3)において、世界的にも例のない特許検索に特化したテストコレクションを構築した[4, 5, 8]。これに関連して、ACL(計算言語学会)の2003年大会で特許コーパス処理に関するワークショップを開催し[3]、そこでは特許文書の検索や解析に関する様々な研究成果が発表された。最近では、特許分類用テストコレクションがWIPOから配布されている[1]。

本稿は、第4回 NTCIR ワークショップ(NTCIR-4)において引き続き遂行している特許検索タスクと、構築中のテストコレクションについて説明する。NTCIR-4は、2003年1月にシステム参加募集を出

した時点から始まり、2004年6月に予定されている成果報告会をもって終了する。現在はタスクが進行中であるため、途中経過の報告と今後の予定について説明する。

## 2 NTCIR-3 特許検索タスクの概要

ある発明が特許として成立し、その権利が消滅する過程において様々な調査が行われる。調査の目的に応じて、性質が異なる特許検索が必要になる。代表的な調査の例として、技術動向の調査や特許庁の審査官が行う実体審査などがある。

調査の目的によって調査範囲(検索対象)や要求される性能(精度と再現率のどちらを重視するか)などが大きく異なるため、汎用的な特許検索コレクションを構築することは容易ではない。

NTCIR-3の特許検索タスクでは技術動向調査を目的としたテストコレクションを構築した<sup>4</sup>。具体的には、ある特定技術に関する新聞記事を検索課題として、1998~1999年に公開された日本語の公開特許公報全文データを検索対象とした(文書数:約70万件、ファイルサイズ:約18GB)。公報全文では規模が大きいなどの理由で扱いづらい研究者を想定し、代替情報として日本語と英語の特許抄録も用意した。

検索課題として日本語の新聞記事を使い、さらに英語、韓国語、中国語(簡体字、繁体字)の翻訳版も用意した。そこで、新聞記事を用いて特許を検索するという「ジャンル横断検索」や、課題と異なる言語で書かれた特許を検索するという「言語横断検索」の性能を評価することが可能である。

前者については(特許検索という側面を考慮しなくても)それまでのテストコレクションにはなかった特性である。その結果、ジャンルごとの索引語重みを考慮する Term Distillation 法[2]の有効性が報告されるなど、特許検索の研究開発促進に対して一定の役割を果たすことができた。

後者の言語横断検索については、各国が海外出願重視の傾向にある現代において今後益々必要とされる技術である。また、日本語と英語の抄録は対訳関係にあるため、対訳語抽出などの言語処理研究にとっても有用性が高い。

## 3 NTCIR-4 特許検索タスクの概要

NTCIR-3では技術動向調査を目的としたのに対して、NTCIR-4では特許の請求項で主張されている権

<sup>1</sup><http://trec.nist.gov/>

<sup>2</sup><http://clef.iei.pi.cnr.it/>

<sup>3</sup><http://research.nii.ac.jp/ntcir/index-ja.html>

<sup>4</sup><http://research.nii.ac.jp/ntcir/permission/perm-ja.html>

利を無効化するための調査(無効資料調査)を目的としている<sup>5</sup>。これは、審査請求があった出願に対して特許庁審査官が行う実体審査(出願が特許権の取得に該当するかどうかを判断するための審査)やサーチャーが他者の権利を無効化するために行う社内での調査などに相当する。

言い換えれば、日本では先願主義を採用しているため、対象となっている発明が既に出願されていたり、既に公知の事実であることを示す資料を検索することが調査の目的である。実際には特許や論文など様々な情報が調査対象になる。しかし、NTCIR-4では公開公報だけを検索対象にした。

NTCIR ワークショップは約1年半を周期として開催される。タスクの計画、データの準備、参加システムの評価などを行うと、1年半は決して長くない。限られた期間で参加者が取り組むことができるタスクに制限すると、中長期的な研究期間を要する手法や技術を評価することは容易ではない。

そこで、参加者が比較的短期間で取り組むことができる(しかし、決して易しくはない)内容をメインタスクとして位置付けつつ、特許検索に関する萌芽的な技術を実践するタスクを実現可能性調査(Feasibility Study: FS)タスクとして並行することにした。

メインタスクは上述した無効資料調査である。FSタスクでは「特許マップの自動作成」を課題とした。FSタスクの目的は、タスクとして実現できるかどうかの見極めや問題点の洗い出しである。結果によっては次回以降の中心となるタスクになる可能性がある。

以下、4章と5章でメインタスクとFSタスクについてそれぞれ説明する。

## 4 メインタスク

### 4.1 概要

メインタスクである無効資料調査の目的は、既存の特許に含まれる請求項を検索課題とし、そこで請求されている権利を無効化できる別の特許を検索することである。検索課題と検索対象のどちらも特許公報であり、類似文書検索と見なすことができる。

本タスクの最終目的は、研究利用のためのテストコレクション(検索課題、文書データ、適合判定)を整備し、公開することである。以下、メインタスクを遂行する手順を示す(図1も参照)。ここで「オーガナイザ」とは、本稿筆者の3名である。課題作成と適合判定を行う「判定者」は、日本知的財産協会(JIPA)知的財産情報検索委員会の委員が担当する。

<sup>5</sup><http://www.slis.tsukuba.ac.jp/fujii/ntcir4/cfp-ja.html>

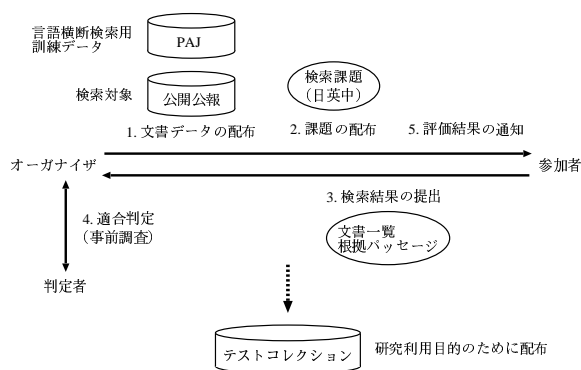


図1: メインタスクの概要

1. 文書データの配布(オーガナイザ → 参加者)  
実際の無効資料調査では過去5年以上の特許を調査することが一般的であるため、NTCIR-3よりも公報データの年数を増やした。NTCIR-4開始当初、参加者には1993~2002年に公開された公開特許公報全文データをDVDで配布した。しかし、公報10年分では検索対象の規模が大きすぎることを主な理由として、1993~1997年の5年分だけをタスクに使用することにした。課題が配布されるまでの期間、参加者は文書データを用いて索引付けや予備実験などを行うことができる。
2. 課題の作成と配布(オーガナイザ → 参加者)  
課題となる特許は、1993~2002年に公開された公報から選択した。さらに、言語横断検索を行うために、英語と簡体中国語に翻訳した課題も用意した。
3. 検索結果の提出(参加者 → オーガナイザ)  
参加者(チーム)は、与えられた課題に対応する特許を検索し、その一覧を提出する。公報は非常に長いので、検索された文書一覧を表示するだけでなく、文書中のどの部分が判断の根拠になっているのかをユーザに提示できることが好ましい。そこで、文書一覧に加えて、各文書において適合する段落(根拠パッセージ)を合わせて提出する。
4. 検索結果の評価(オーガナイザ ↔ 判定者)  
提出された結果を用いて適合文書の候補を収集し、それらに対して適合判定を行う(ブリーディング法[6])。ただし、この作業とは独立に、判定者が事前調査を行って適合文書を集めておく。

## 5. 評価結果の通知（オーガナイザ → 参加者）

完成した適合文書群を用いて、提出された結果を評価し、各参加者に結果を通知する。

なお、オーガナイザ、判定者、参加者が互いに手順を確認したり、タスクの問題点を発見するために、ドライラン（予備試験）とフォーマルラン（本試験）に分けて上記の手順を2回繰り返す。ドライランは既に終了し、フォーマルランの結果提出を2003年11月下旬に締め切った。フォーマルランの評価結果は2004年2月下旬に通知する予定である。

## 4.2 文書データ

検索対象となる文書データは、1993～1997年に公開された日本語の公開特許公報全文データである。これらは出願から18ヶ月たって公開される特許全文データで、特許庁から公開されているものと同じである。ただし、図表情報は含まない。文書数は約170万件、ファイルサイズは約40GBである。既存の情報検索用テストコレクションに比べて、文書数、ファイルサイズともかなり大きい。

各公報には公開番号に準拠して文書IDを付与する。また、パッセージ検索の結果を提出するため、全参加者がパッセージの単位を共有しなければ評価が困難になる。そこで、墨付括弧でくられた段落記号（【0002】や【従来の技術】）を基準として（文章内容の論理構造とは無関係に）パッセージを機械的に認定する。これらの情報を示すタグを自動的に挿入する公式ツールを参加者に配布し、参加者が個別にデータの前処理をできるようにした。

また、主に言語横断検索向けの訓練データとして「日本国公開特許英文抄録データPAJ」を配布した。これは、1993～2003年に公開された公報の「要約」を英語に翻訳したコーパスである（要約を加筆・編集した後で翻訳している場合がある）。

## 4.3 検索課題

### 4.3.1 課題作成方法の概要

検索課題は、検索対象と同じく公開公報に基づいて作成した。ただし、5年分ではなく、1993～2002年の10年分を選択した。

課題の作成には、日本知的財産協会（JIPA）知的財産情報検索委員会の協力を得ることができた。当委員会の委員は民間企業知財部のサーチャーで、特許検索の専門家である。各委員（判定者）が自分の

専門分野に関する課題を作成し、その課題に対する適合判定も担当する。

ただし、どのような公報でも課題にできる訳ではない。適合文書が確実に存在する課題を作るために、既に審査請求がされ、なおかつ特許庁審査官が拒絶した出願を選んだ。そうすることで、最低でも審査官が提示した引例（拒絶する根拠となる特許）を適合文書とすることができる。

さらに、適合文書が多いことも重要な要件である。課題あたりの適合文書数が極端に少ないと、評価結果が不安定になるということが過去のテストコレクション構築研究で知られているからである。

既存のテストコレクションでは、1つの課題に対して、平均100以上の適合文書が存在することは珍しくない。それに対して、判定者が行った予備調査では、無効資料調査において適合文書が10件を超える課題を選定することが困難であった。

適合文書が少ない場合は、課題数を増やすなどの対処が必要になる。例えば、質問応答システムの評価では、1つの質問に対して正解が1つか数件であるため、数百件の質問を用意することがある[7]。

しかし、主に人的コストの制約によって30件程度の課題を作ることが限度であった。無効資料調査における課題の作成や適合判定には、特許検索や対象技術についての高度な専門知識が要求されるため高価である。

そこで、2種類の課題を作成することでこの問題に対処した。まず、課題となる公報の選定と適合文書の判定を判定者が精緻に行う「メイン課題」である。この課題は作成や判定が高価で件数が限られる。もう1つは、特許庁審査官が拒絶した引例だけを適合文書として利用する「追加課題」である。公報の選定は判定者が行うものの、適合判定が不要であるため、安価で多数の課題を作ることができる。

これら2種類の課題を個別もしくは併用して評価に用いることで、システム間の評価結果がどのように変動するかを調査し知見を得ることは、今後の特許検索用テストコレクションの構築にとって重要なことである。

### 4.3.2 メイン課題

人手で判定することを前提にした「メイン課題」では、特許庁審査官が提示した引例に加えて、判定者が事前に調査（通常業務と同じ環境における検索活動）を行い、適合文書がある程度の件数見つかった公報を中心に選定した。

検索課題の抜粋を図2に示す。この課題では請求項1が対象となっている。対象請求項には<CLAIM>

が付与されており、この部分を検索システムの入力とすることが原則である。

実際の無効資料調査では、請求項を漠然と眺めるのではなく、発明や技術のポイントを考慮して構成要素に分割し、なるべく多くの(主要な)構成要素に合致するような特許を適合文書と判断する。

そこで、参加者が人間の調査方法を模倣したシステムを実装しやすいように、構成要素に関するタグ情報(<COMP>)が判定者によって付与されている。ただし、このタグ情報は使用しなくてもよい。

さらに、なるべく実際の調査を模倣するために、課題特許の出願日を基準日とし、基準日よりも前に公開された公報だけを検索対象にした。ただし、優先権主張制度に基づいて出願された場合は基準日の認定が困難な場合がある。そこで、基準日を統一するためのタグ(<FDATE>)が付与されている。

上記の手順によって、ドライランで7件、フォーマルランで34件の課題を作成し、評価に利用した。さらに、英語と簡体中国語の翻訳も用意した。最終的なテストコレクションには他の外国語翻訳を含める可能性がある。

#### 4.3.3 追加課題

追加課題は、判定者による判定を行わず、特許庁審査官が拒絶の根拠として使用した引例特許だけを適合文書として評価に利用する。ただし、対象の請求項は何番か、引例が1993～1997年に公開された公報データに含まれるかどうか、出願から拒絶までの期間中に内容が補正されていないかなどの種々の簡単なチェックを行い、適切な課題を選定した。課題の形式はメイン課題(図2)とほぼ同じである。しかし、メイン課題とは以下の点が異なる。

- 課題作成や適合判定のコストが安価であり、比較的大量に用意できる。
- 適合文書の件数が非常に少ない。1件のものが大半である。
- 人手コストをかけないという理由から、請求項を構成要素に分割しておらず、<COMP>タグは含まれない。ただし、NTCIR-4ワークショップ後に配布するテストコレクションでは構成要素の分割情報を含める可能性がある。
- 外国語の翻訳は用意しない。

69件の追加課題を作成し、参加者に配布した。今後必要に応じて課題数を増やし、最終的なテストコレクションに含めることも検討している。

#### 4.4 参加者が行う作業

参加者が行う作業の基本は、課題ごとに検索された文書一覧と文書ごとの根拠パッセージ一覧をオーガナイザに提出することである。

そのためには、配布された検索課題をシステムに入力可能な検索質問(クエリ)に変換する必要がある。クエリの形式として論理式やキーワード列が一般的である。ここで、クエリ作成手法について以下に示す選択肢が与えられる。

- 全自動: 課題ファイルからクエリを機械的に作成する。
- 半自動: 課題ファイルからクエリを作成する段階で人間の補助が必要になる。参加者自身が手作業を行わなくても、図2の<COMP>タグを使用した場合は、事実上、人手による作業が必要となるため、半自動によるクエリ作成と見なす。
- 手動: 課題ファイルからクエリを作成する段階で、人間が中心になって作業を行う。

また、課題ファイルは、対象となる請求項以外にも様々な情報を含んでいるため、使用する情報にも選択肢がある。

- 対象請求項だけを用いてクエリを作成する。
- 対象請求項以外の情報も用いてクエリを作成する。例えば、国際特許分類(IPC)や発明内容を詳細に記述した項目を用いる手法が考えられる。

そこで、課題ファイルからクエリを作成する段階において、既存の検索問題とは異なる様々な戦略を立てる余地がある。どのような手法が有効であるかは、評価結果によって明らかになるであろう。

提出内容は、既存のシステム参加型ワークショップと同様に、配布された課題ごとに、検索された文書を適合度に基づいて順位付けした上位1000件までの一覧である。それに加えて、NTCIR-4特許検索タスクでは、各文書についてパッセージを「その文書が適合と判断する根拠になる」という基準で順位付けし、提出する。そこで、文書検索に加えてパッセージ検索が必要になる。

また、基準日(図2の<FDATE>)よりも前(基準日を含まない点に注意)に公開された公報だけが適合文書の候補となる。仮に基準日かそれ以降に公開された公報を提出しても評価には使用されない。そこで、テキスト検索技術に加えて、複数のフィールドによる絞り込み検索技術が要求される。

なお、参加者はクエリ作成や検索に使用したデータや手法などを申告することになっている。

```

<TOPIC>
<NUM>008</NUM>
<LANG>JA</LANG>
<PURPOSE>invalidity search</PURPOSE>
<FDATE>19960527</FDATE>
<DOC>
<DOCNO>PATENT-JA-UPA-1997-318650</DOCNO>
<TEXT>
<SDO BIJ><DP N=0001><RTI ID=000001 HE=150 WI=170 LX=0200 LY=0300>(19)【発行国】日本国特許庁( J P )
(12)【公報種別】公開特許公報( A )
(11)【公開番号】特開平 9 - 3 1 8 6 5 0
(43)【公開日】平成 9 年( 1 9 9 7 ) 1 2 月 1 2 日
(54)【発明の名称】センサ装置及びその製造方法
(51)【国際特許分類第 6 版】
    G01P 15/09
    G01L 9/08
... ( 中略 ) ...
</RTI></SDO><SDO ABJ><TXF FR=0001 HE=045 WI=080 LX=0200 LY=1800>(57)【要約】
【課題】 部品点数を少なくしてセンサを薄型化、小型化することができるセンサ装置及びその製造方法を提供し、更には、センサを安価に製造できるセンサ装置及びその製造方法を提供する。
【解決手段】 箱状の成形基台 2 に開放された凹所 1 を形成し、前記凹所 1 の内面を含む成形基台 2 の表面に所定のパターンの導電膜を形成し、前記凹所 1 にセンサ用素子 4 を接合し、成形基台 2 を蓋 5 で密閉した。
<EMI ID=000002 HE=100 WI=080 LX=1100 LY=1800></SDO><SDO CLJ><DP N=0002>
<TXF FR=0001 HE=250 WI=080 LX=0200 LY=0300>
<PASSAGE>
<PNUM>PATENT-JA-UPA-1997-318650-0</PNUM>
【特許請求の範囲】
<CLAIM>
【請求項 1】 <COMP><CNUM>1</CNUM>箱状の成形基台に開放された凹所を形成し、</COMP><COMP><CNUM>2</CNUM>前記凹所の内面を含む成形基台の表面に所定のパターンの導電膜を形成し、</COMP><COMP><CNUM>3</CNUM>前記凹所にセンサ用素子を接合し、</COMP><COMP><CNUM>4</CNUM>成形基台を蓋で密閉してなる</COMP><COMP><CNUM>5</CNUM>ことを特徴とするセンサ装置。
</CLAIM>
</PASSAGE>
<PASSAGE>
<PNUM>PATENT-JA-UPA-1997-318650-1</PNUM>
【請求項 2】 センサ用素子が加速度センサ用素子であることを特徴とする請求項 1 記載のセンサ装置。
</PASSAGE>

```

図 2: 検索課題の抜粋 ( 課題番号: 008 )

以上をまとめると、参加者が行う作業手順は次のように一般化することができる。

1. 検索対象文書データの索引付け
2. 課題ファイルからのクエリ作成 ( 言語横断検索では翻訳も必要 )
3. 文書検索 ( 日付による絞り込みが必要 )
4. パッセージ検索

これらの手順を全て行うことが困難な参加者のために、上記 1 と 3 を行う検索システムをオーガナイザが構築し、オンラインで利用するための API を参加者に提供した。そこで、クエリの作成や翻訳、検索された文書からのパッセージ検索を行うだけでもタスクに参加することができる。

また、複数のチームが基盤となる検索システムを共有することには、それ以外の要素技術を比較評価することが容易になるという利点がある。

#### 4.5 評価方法

参加者が提出した検索結果は、文書単位とパッセージ単位の 2 つの観点で個別に評価する。両者を統合した評価手法は今後の検討課題である。

文書単位の評価は、原理的には、既存の評価型ワークショップで行う手法と同じである。参加者が提出した検索結果を用いて適合文書の候補を収集し、それらの候補に対して判定者が適否を判定する。さらに、特許庁審査官が拒絶の根拠として実際に使った引例や判定者が事前に取得した公報も適合文書の候補とする。適合判定の基準は、課題の請求項で主張され

ている権利を無効化できるかどうかである。具体的には、以下の基準に基づいて多段階に判定する。

- A: 単独で無効化できると判定された公報。
- B: それだけでは無効化できないため、他の公報と組み合わせて根拠を補強する必要があると判定された公報。ただし、当該公報が根拠の大半を占めていることが条件である。また、組み合わせる相手が実在するかどうかは問わない。
- C: 内容を吟味した上で不適合と判定された公報。
- D: 発明の名称だけで不適合と判定された公報。

最も厳しい評価では A だけを正解と見なし、評価基準を緩める場合は B も正解と見なす。いずれの場合も、C と D は不正解として扱う。

評価尺度には通常用いられる平均精度 ( Mean Average Precision: MAP ) を用いる予定である。しかし、通常の特許検索では多数の文書 ( 例えば、300 件 ) を調査することが一般的である。そこで、順位の微妙な差異は問題にならないことがある。

それに対して、MAP は適合文書の順位が変動すると値が大きく変わることがある。オーガナイザが統一的な尺度を用いて全参加者の結果を評価する一方で、各参加者には成果報告会において独自の尺度による評価も推奨してる。こうした一連の活動を通して、どのような評価尺度が適しているのか今後検討する必要がある。

パッセージ単位の評価を行うために、適合と判定された文書 ( 上記の A と B ) に対して、適合と判断する根拠になるパッセージを判定者が網羅的に特定する。類似したパッセージが複数箇所に含まれる場合は、それらを全て特定する。

また、複数のパッセージを組み合わせないと根拠にならない場合は、可能な組み合わせ方を全て特定する。そこで、単独では根拠にならなくても複数組み合わせることによって単独の根拠パッセージと同じだけの情報を持つ「根拠グループ」を考慮した評価が可能になる。言い換えれば「組み合わせ適合」という新しい概念を情報検索の評価に取り入れることが可能になる。

従来の情報検索に関する評価では、適合文書が全て独立に有用な情報であるという仮定があった。それに対して、単体では不十分な文書でも、組み合わせ次第では単一の適合情報と等価なものとして扱うという点に特殊性がある。

パッセージ検索の評価尺度は確立されていない。現在は、根拠パッセージに到達するまで、もしくは根拠グループ内のパッセージを全て網羅するまでの平均探索長を用いる予定である。

## 5 FS タスク

FS タスクでは特許マップの自動作成を課題とする。特許マップとは、既存の特許を一定の基準で分析し、その結果を可視化したものである。特許マップを利用する目的によって、分析の範囲や方法、その形態、可視化の方法は様々である。NTCIR-4 では、特定の技術分野や製品などのテーマに関連する特許群を 2 次元のマトリクスで表現することを目的とする。

図 3 は「光ディスク」をテーマにした特許マップの例である。ここでは、横軸と縦軸がそれぞれ「開発課題」と「達成手段」に対応しており、各セルに該当する特許 ID ( P1 や P2 など ) が表示されている。

		開発課題		
		高密度記録	消去特性の向上	書換回数 の向上
達成手段	書換回数の管理	P1, P2		P3
	記録位置のシフト		P4, P5	P6

図 3: 「光ディスク」をテーマにした特許マップの例

実際に特許マップを作成する場面では、テーマや分類の軸 ( 開発課題や達成手段など ) を設定した上で、それぞれの軸に対応する具体的な項目 ( 高密度記録など ) を決定し、該当する特許をマトリクスのセルに分類する。

ここで、該当する特許をある程度絞り込むための検索が必要になる。しかし、特許マップの作成に焦点を絞るため、NTCIR-3 特許検索テストコレクションを使うことにした。当コレクションでは、ある技術に関する適合特許が既に特定されているからである。

以下、参加者に配布される課題と参加者が提出する内容を NTCIR-3 のテストコレクションとの関係に注意しながら示す。

- 課題
  - － テーマ  
NTCIR-3 テストコレクション 検索課題のテーマ ( 「ガソリン直噴エンジン」など )
  - － 縦軸と横軸  
図 3 における「開発課題」や「達成手段」など ( ただし、課題によって異なる )
  - － 分類対象となる特許コーパス  
NTCIR-3 のテストコレクションで適合と判定された特許
- 提出内容  
図 3 における「高密度記録」や「記録位置のシフト」などの分類項目と分類対象となる特許の対応関係

NTCIR-3 テストコレクションに含まれる 31 課題のうち、適合文書数が 100 件程度存在するものを 6 件選択し、FS タスクの課題とした。NTCIR-3 当時の課題作成者が FS タスクにおいても改めて課題を作成し、また模範解答となる特許マップを作成した。

特許マップ作成の自動化に関する一般的なモデルは確立されていない。しかし、分類項目となるキーワードを公報から複数抽出し、それらの項目に基づいて対象公報を分類する技術が必要になるだろう。

NTCIR-3 テストコレクションでは、検索課題として日本語、英語、韓国語、中国語(簡体字、繁体字)があり、どれを使用してもよい。また、公開公報、JAPIO 抄録(和文)、PAJ(英文)という 3 種類のコーパスがある。参加者はどのコーパスを利用してよい。そこで、多言語・言語横断の特許マップ作成も可能である。

提出された特許マップの評価は定量的な方法ではなく、模範解答のマップと比較することで定行的に行い、必要に応じて添削することを予定している。

FS タスクで作成されるテストコレクションは、検索と分析(分類)というタスクが有機的に結合している点に特殊性がある。

全自動の手法でなくても、半自動の手法や人間中心型の作成手法による提出も認めている。

## 6 おわりに

NTCIR-4 特許検索タスクでは、無効資料調査を中心的な検索タスクとし、さらに特許マップの作成を試験的なタスクとして並行している。今回の特許検索タスクを一般的な情報検索の問題と捉えた場合、最終的に完成するテストコレクションは、以下のような研究課題を探索するために有効である。

- 単件ごとの文書長が非常に長く、大規模なコレクションに対する検索
- 類似文書検索(公報で公報を検索する)
- マルチトピック検索(検索課題の請求項が複数の要素で構成されている)
- パッセージ検索
- 適合情報間の依存関係(組み合わせ適合)を考慮した検索
- 構造やメタデータを持つ文書の検索
- 分析や分類と結合した検索(FS タスク)

(特許検索そのものには関心がなくても)これらのいずれかに興味がある研究者のために NTCIR-4 特許検索テストコレクションが役立てば幸いである。

## 謝辞

日本知的財産協会知的財産情報検索委員会の関係各位には、NTCIR-3 に引き続き、NTCIR-4 特許検索タスクの遂行においても多大なご支援を頂いています。この場を借りて深謝いたします。

## 参考文献

- [1] C. J. Fall, A. Töröcsvári, K. Benzineb, and G. Karetka. Automated categorization in the international patent classification. *ACM SIGIR FORUM*, Vol. 37, No. 1, pp. 10–25, 2003.
- [2] Hideo Itoh, Hiroko Mano, and Yasushi Ogawa. Term distillation for cross-DB retrieval. In *Proceedings of the Third NTCIR Workshop on Research in Information Retrieval, Automatic Text Summarization and Question Answering*, 2003.
- [3] Makoto Iwayama and Atsushi Fujii, editors. *Proceedings of the ACL-03 Workshop on Patent Corpus Processing*, 2003.
- [4] Makoto Iwayama, Atsushi Fujii, Noriko Kando, and Yuzo Marukawa. An empirical study on retrieval models for different document genres: Patents and newspaper articles. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 251–258, 2003.
- [5] Makoto Iwayama, Atsushi Fujii, Noriko Kando, and Akihiko Takano. Overview of patent retrieval task at NTCIR-3. In *Proceedings of the Third NTCIR Workshop on Research in Information Retrieval, Automatic Text Summarization and Question Answering*, 2003.
- [6] Ellen M. Voorhees. Variations in relevance judgments and the measurement of retrieval effectiveness. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 315–323, 1998.
- [7] Ellen M. Voorhees and Dawn M. Tice. Building a question answering test collection. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 200–207, 2000.
- [8] 岩山真, 藤井敦, 高野明彦, 神門典子. 特許コーパスを用いた検索タスクの提案. 情報処理学会研究報告, 2001-FI-63, pp. 49–56, 2001.