

Chapter 4

Challenges in Patent Information Retrieval



Makoto Iwayama, Atsushi Fujii, and Hidetsugu Nanba

Abstract We organized tasks on patent information retrieval during the decade from NTCIR-3 to NTCIR-8. All of the tasks were ones that reflected real needs of professional patent searchers and used large numbers of patent documents. This chapter describes the designs of the tasks, the details of the test collections, and the challenges addressed in the research field of patent information retrieval.

4.1 Introduction

A patent for an invention is a grant for the inventor to exclusively exploit the invention in the limited term in return to disclosing it to the public. The invention is described in a document called a patent application (also called a patent specification or an application document), which is composed of an abstract and sections describing the scope of the invention (the claims), the problems to be solved, the embodiments of the invention, etc. The patent application is filed with the patent office. The date of filing is called the filing date or application date. After the filing, the patent office examines the patent application, and if the invention is judged to be novel, in other words, one which has no prior art, a patent is granted for it.

As the economy grows worldwide, the number of patent applications and grants has also grown. The World Intellectual Property Organization (WIPO) announced

M. Iwayama (✉)

Hitachi Ltd. Research & Development Group, 1-280, Higashi-Koigakubo,
Kokubunji-shi, Tokyo 185-8601, Japan
e-mail: makoto.iwayama.nw@hitachi.com

A. Fujii

School of Computing, Tokyo Institute of Technology, W8-62, 2-21-1 Ookayama,
Meguro-ku, Tokyo 152-8552, Japan
e-mail: fujii@cs.titech.ac.jp

H. Nanba

Department of Industrial and Systems Engineering, Chuo University,
1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan
e-mail: nanba@kc.chuo-u.ac.jp

© The Author(s) 2021

T. Sakai et al. (eds.), *Evaluating Information Retrieval and Access Tasks*,
The Information Retrieval Series 43,
https://doi.org/10.1007/978-981-15-5554-1_4

that the number of patent applications in 2017 had exceeded three million. In Japan, about three-hundred-thousand patent applications are filed every year. Since patent applications are highly technical and their length tends to be long, the task of searching for patent applications poses many issues in relation to information retrieval; similar situations are searching for technical papers or searching for legal documents.

In this chapter, we introduce the challenges aimed at addressing the issues of patent information retrieval.¹ These challenges were formulated as tasks performed in NTCIR workshops from 2001 to 2010. The NTCIR tasks were designed on the basis of actual patent-related work involving a large number of patent applications. The remainder of this chapter is organized as follows. Section 4.2 briefly introduces the NTCIR tasks. Section 4.3 describes the tasks in detail, including the search topics, document collections, submissions, relevance judgements, evaluation measures, and participants. Finally, Sect. 4.4 summarizes NTCIR's contributions to the research activities on patent information retrieval.

4.2 Overview of NTCIR Tasks

4.2.1 *Technology Survey*

Managers, researchers, and developers often want to know whether there are existing inventions related to the products they are planning to develop. This situation is similar to when researchers survey research papers before embarking on new research.

To satisfy this information need, they have to conduct a “technology survey” that involves searching for relevant patent applications published so far. Here the query might not be described in patent-specific terms, because the searcher is not always familiar with the procedure for searching for patents. Moreover, the notion of relevance is not patent-specific. Patent applications are treated like technological articles such as research papers.

4.2.2 *Invalidity Search*

After inventing a new method, device, material, etc., the inventor describes the invention in a patent application and sends it to the patent office. The patent office then examines the application to see if there is prior art which invalidates the invention by searching for patent applications filed before the filing date. This is called a “invalidity search” or “prior art search”. Invalidity searches are also conducted by applicants themselves, because they should be confident of their inventions being granted patents before they make their applications.

¹Readers who are interested in patent machine translation can refer to Chap. 7.

The invalidity search is patent-specific work. A searcher should be able to understand the components of an invention in accordance with the claims described in the application. Relevance is assessed based on the novelty or the invalidity of the invention. The searcher compares each component of the invention with portions of each retrieved document to see if they describe the same invention. If there is no prior art which can invalidate the novelty of the invention, a patent is granted; otherwise, the application is rejected. In most cases of rejection, several instances of prior art are cited, each of which corresponds to a component of the described invention.

4.2.3 Classification

Classification codes are extensively used when searching to narrow down the relevant applications. The patent office assigns each patent application appropriate classification codes before it is published. Human experts have to expend much effort to make this assignment, and for this reason, a (semi-)automatic method is desired.

The most popular classification codes for patents are the International Patent Classification (IPC) codes which are used worldwide. The Japan Patent Office (JPO) additionally uses and maintains a list of F-terms (File forming terms). F-terms are facet-oriented classification codes, and a patent application is classified from a variety of facets (viewpoints) such as objective, application, structure, purpose, and means.

In NTCIR, patent applications are automatically classified with F-terms in accordance with the behavior of human experts who perform their classification work in two steps. The first step is the theme (topic) classification, assigning a patent application to technological themes. Each theme corresponds to a group of IPC “sub-classes”. The number of themes is about 2,500. The second step is F-term classification, i.e., assigning F-terms to an application that has already been assigned themes in the first step. Although the total number of F-terms is huge, over 300,000, the number of F-terms within each theme is relatively small, about 130 on average.

4.2.4 Mining

For a researcher in a field with high industrial relevance, analyzing research papers and patents has become an important aspect of assessing the scope of their field. The JPO creates patent application technical trend surveys for fields in which the development of technologies is expected, or fields to which social attention is being paid. However, it is costly and quite time-consuming.

In NTCIR, we aimed to construct a technical trend map from research papers and patents in a specific field. For the construction of the map, we focused on the elemental (underlying) technologies used in a particular field and their effects. Knowledge of the history and effects of the elemental technologies used in a particular field is

important for grasping the outline of technical trends in the field. Therefore, we designed the task to extract elemental technologies and their effects from research papers and patents.

4.3 Outline of the NTCIR Tasks

4.3.1 Technology Survey Task: NTCIR-3

Table 4.1 summarizes the test collections of the NTCIR tasks for patent retrieval.

A technology survey task was performed at only NTCIR-3 (Iwayama et al. 2003). Since this task was our first attempt to handle a practical number of patent applications, we designed the task to be as close as possible to the ad hoc retrieval tasks in TREC, except that the targeting documents were patent applications.

To launch the task, we obtained the cooperation of members of the Japan Intellectual Property Association (JIPA), who are experts in patent searches. Each JIPA member belongs to the intellectual property division in the company he or she works for. We collaborated with them in designing our tasks, constructing search topics, collecting relevant documents, evaluating the submitted results, and many other ways. Our collaboration continued through to NTCIR-4, and this was a major reason for the success of our challenges.

4.3.1.1 Search Topics

The technology survey task assumed a situation where a searcher is interested in a technology, for example, a “blue light-emitting diode”, described in a newspaper

Table 4.1 NTCIR test collections for patent retrieval

Task		NTCIR-3	NTCIR-4		NTCIR-5	NTCIR-6	
		Technology survey	Invalidity search				
			Main	Additional		Main	English
Search topics		31(+6) ja,en,zh-CN,zh-TW	34(+7) ja,en,zh-CN	69 ja	1189 ja	1685 ja	2221(+1000) en
Document collections	Patent applications	1998–1999 ja	1993–1997 ja		1993–2002 ja		1993–2000 en
	Abstracts	1995–1999 ja,en	1993–1997 en		1993–2002 en		–
Relevance judgments		Manual		Citation			

article. The JIPA members constructed 31 search topics from newspaper articles which, in most cases, were selected from the topics they were working on in their daily jobs. Each search topic contained a title, headline, and text of the article that triggered the request for information, a description and a narrative of the topic, a set of concepts (keywords) related to the topic, and a supplement with more information about the relevance. All the search topics were translated into English, Korean, and Chinese (simplified and traditional).

4.3.1.2 Document Collections

The main documents for the retrieval were Japanese full texts of (unexamined) patent applications published in 1998 and 1999. The number of documents was 697,262. We also released abstracts of patent applications published over the 1995–1999 period, in Japanese and in English. The English abstracts were translations from the Japanese ones. The number of documents was 1,706,154 for the Japanese abstracts and 1,701,339 for the English abstracts. Some of the Japanese abstracts did not have corresponding English abstracts.

4.3.1.3 Submissions

Each participant submitted at least one run that used only the newspaper articles and supplements on the given search topics. In addition, we recommended that they submit ad hoc runs that used the descriptions and the narratives. For each search topic, a ranked list of at most 1000 patent applications was submitted in decreasing order of relevance score.

4.3.1.4 Relevance Judgments

The relevance of the technology survey is not patent-specific; that is, it is not based on the novelty of the invention, but rather on the relatedness of the search topic to the patent application.

The relevance was assessed by JIPA members in two steps. In the first step, the JIPA member who created the topic collected relevant documents on the topic before its release. Here although the members were allowed to use any search tools, almost all of them used Boolean ones, despite the fact that the organizers had provided a rank-based search system. In the second step, after the participants submitted their results for a topic, the JIPA member who created the topic judged the relevance of the unseen documents in the pool collected from the top-ranked submitted documents.

4.3.1.5 Evaluation

The submitted runs were evaluated by comparing recall/precision trade-off curves and values of mean average precision (MAP).

4.3.1.6 Participants

Eight groups submitted 36 runs. The top-performing run was from Ricoh (Itoh et al. 2003). They focused on re-weighting terms based on their statistics in the different collections (patent applications vs. newspaper articles). For example, the query term “president” in a newspaper article might not be effective for retrieving relevant patent applications. However the inverse document frequency (IDF)-based weighting gives this term a large weight, because it occurs rarely in patent applications. Their approach, called “term distillation”, involved multiplying the weights in the query (i.e., newspaper articles) and the target documents (i.e., patent applications) to select effective terms from a query newspaper article.

4.3.2 *Invalidity Search Task: NTCIR-4, NTCIR-5, and NTCIR-6*

Having gained experience in the technology survey task at NTCIR-3, we moved on to invalidity search, which is a patent-specific search. Invalidity search tasks were performed in NTCIR-4 (Fujii et al. 2004), NTCIR-5 (Fujii et al. 2005), and NTCIR-6 (Fujii et al. 2007b).

4.3.2.1 Search Topics

Invalidity searches are searches, for instances, of prior art that could invalidate a patent application. Here, the patent application itself becomes a search topic. As search topics in NTCIR-4, JIPA members selected 34 Japanese patent applications that had been rejected by JPO. We called this set “NTCIR-4 main topics”.

Regarding the NTCIR4 main topics, relevant documents were thoroughly collected by the JIPA members (see Sect. 4.3.2.4 for the details of this collection procedure). However, we found that the number of relevant documents in invalidity searches was small compared with the existing test collections for information retrieval. Consequently, evaluations made on a small number of topics could potentially be inaccurate. To increase the number of search topics, JIPA members selected an additional 69 search topics from other rejected patent applications. Here, relevant documents were only the citations reported by JPO. We called this set “NTCIR-4 additional topics”. Note that the major difference between these two sets relates to

the completeness of relevant documents. We will discuss the effect of this issue on the evaluation in Sect. 4.3.2.6.

The NTCIR-4 main topics set had two more resources with which to make additional evaluations. First, each claim was translated into English and simplified Chinese for evaluating cross-language retrieval. Second, each Japanese claim had annotations for the components of the invention. Components were, for example, parts of a machine or substances of a chemical compound. Some participants used component information in the task.

In both NTCIR-5 and NTCIR-6, the organizers increased the number of search topics by following the same method used to create NTCIR-4 additional topics. Accordingly, the relevant documents for these topics were only citations. The number of search topics was 1,189 in NTCIR-5 and 1,685 in NTCIR-6. We called these topic sets “NTCIR-5 main topics” and “NTCIR-6 main topics”.

NTCIR-5 included a passage retrieval task as a sub-task of the invalidity search task. Since patent applications are lengthy, it is useful to point out significant fragments (“passages”) in a relevant application. In the passage retrieval task, a relevant application retrieved from a search topic was given, and the purpose was to identify the relevant passages in the relevant application. We used 378 relevant applications obtained from 34 search topics of NTCIR-4 main topics plus another 6 that had been used in the dry run in NTCIR-4.

NTCIR-6 involved an invalidity search task on English patent applications (called the “English retrieval task”). The design of the task was the same as the Japanese one. Each search topic was a patent application published by the United States Patent and Trademark Office (USPTO) in 2000 or 2001. We collected 3,221 search topics (1,000 for the dry run and 2,221 for the formal run) from those satisfying the two conditions; first, at least 20 citations are listed, and second, at least 90% of the citations are included in the target document collection. These citations were relevant documents.

4.3.2.2 Document Collections

In NTCIR-4, the document collection for the target of searching consisted of 5 years’ worth of Japanese (unexamined) patent applications published from 1993 to 1997. The number of documents totaled 1.7 million. We additionally released English abstracts that were translations of the Japanese abstracts in these applications.

In NTCIR-5 and NTCIR-6, the document collections (both Japanese patent applications and English abstracts) were enlarged to include those published over the 10 year period from 1993 to 2002. The number of documents in each collection was 3.5 million.

In the English retrieval task at NTCIR-6, the document collection was patent applications published from USPTO over the period from 1993 to 2000. The number of documents totaled about 1 million.

4.3.2.3 Submissions

Although the full texts of the patent applications were provided as search topics, each participant was requested to submit a result which only used the claims and the filing dates in the search topics. Participants could submit additional results, in which they could use any information in the search topics. The number of documents retrieved for each search topic was 1,000 at maximum, and these documents were submitted in decreasing order of relevance score.

To assess effectiveness across different sets of search topics, each participant was requested to submit a set of results from all the Japanese main topics released so far. For example, in NTCIR-6, each participant had to submit runs using NTCIR-4 main topics and NTCIR-5 main topics in addition to NTCIR-6 main topics.

In the passage retrieval task in NTCIR-5, each participant was requested to sort all passages in each of the given relevant applications according to the degree to which a passage provided grounds to judge if the application was relevant.

4.3.2.4 Relevance Judgments

In invalidity searches, the most reliable relevant documents are ones cited by the patent office when rejecting patent applications. However, we were not confident that using only citations would be enough to evaluate the participating systems from the standpoint of recall. Therefore, in NTCIR-4, we exhaustively collected relevant documents by performing the same two steps that were used in the technology survey task of NTCIR-3.

First, the JIPA members who created the search topics (NTCIR-4 main topics) performed manual searches to collect as many relevant documents as possible. Citations from the topic applications were included among the relevant documents. We allowed the JIPA members to use any system or resource to find relevant documents. In this way, we would obtain a relevant document set under the circumstances of their daily patent searches. Most members used Boolean searches, which to this day remains the most popular method used in invalidity searches. Second, after the participants submitted their runs, the JIPA members judged the relevance for the unseen documents in the pool collected from the top-ranked documents in each run. Here, one promising result was that the participating systems could find a relatively large number of relevant documents which were neither citations nor relevant documents found by the JIPA members in the first step.

In NTCIR-5 and NTCIR-6, we used only citations as relevant documents, mainly because we could not cooperate with expert searchers.

Relevance was automatically graded as relevant, partially relevant, or irrelevant. A document that could solely be used to reject an application was regarded as relevant. A document that could be used with other documents to reject an application was regarded as partially relevant. Other documents were regarded as irrelevant.

In NTCIR-6, we tried an alternative definition of the relevance grade, one based on the observation that if a search topic and its relevant application have the same IPC

codes, systems could easily retrieve the relevant application by using IPC codes as filters. We divided the relevant documents into three classes according to the number of shared IPC codes between the topic and a relevant document and compared the submitted runs on the basis of the classes (Fujii et al. 2007b).

In the passage retrieval task of NTCIR-5, we reused the search topics in NTCIR-4 and all the relevant passages had been collected in NTCIR-4 by JIPA members. Relevance was graded as follows. If a single passage could be grounds to judge the target document as relevant or partially relevant, this passage was judged to be relevant. If a group of passages could be grounds, each passage in the group was judged as partially relevant. Otherwise, the passage was judged as irrelevant.

4.3.2.5 Evaluation

We used MAP for the evaluation measure in all the invalidity search tasks. In the passage retrieval task, we additionally used the averaged passage rank at which an assessor obtains sufficient grounds to judge whether a target document is relevant or partially relevant, when the assessor checked the passages in the top-ranked to bottom-ranked target documents.

4.3.2.6 Participants

Eight groups participated in the invalidity search tasks of NTCIR-4, ten in NTCIR-5, and five in NTCIR-6. The passage retrieval task of NTCIR-5 had four groups, while the English retrieval task of NTCIR-6 had five groups. In this section, we introduce only those groups who participated in most of the main tasks, i.e., document-level invalidity searches using Japanese topics.

Hitachi submitted runs to all of the invalidity search tasks (Mase et al. 2004, 2005; Mase and Iwayama 2007). From NTCIR-4 to NTCIR-6, they tried various methods, for example, using stop words, filtering by IPC codes, term re-weighting, or using the claim's structure. The methods were composed of two-step searches. The first step was a recall-oriented search, and the second step was a re-ranking of the documents retrieved by the first step to improve precision.

NTT Data participated in NTCIR-4 (Konishi et al. 2004) and NTCIR-5 (Konishi 2005). They expanded the query terms with keywords selected from the “detailed descriptions of the invention” (“embodiments”) section. First, they decomposed a topic claim into components of the invention by using pattern-matching rules. Next, they identified descriptions that explain each component by using another set of pattern-matching rules. Lastly, they added keywords in the descriptions to the query terms.

The University of Tsukuba participated in all of the invalidity search tasks (Fujii and Ishikawa 2004, 2005; Fujii 2007). They automatically decomposed a topic claim into components and searched the components independently. Then, they integrated the results. Query terms were also extracted from related passages automatically

identified in the topic document. Retrieved documents which did not share the IPC codes of the query application were filtered out. They observed that the IPC filtering was more effective in NTCIR-5 main topics than in NTCIR-4 main topics (Fujii and Ishikawa 2005; Fujii 2007). This difference might have been due to the nature of the relevance of the two sets. The relevant documents for NTCIR-4 main topics were manually collected, while those for NTCIR-5 main topics were only citations. If we imagine that searches at a patent office often rely on metadata (IPC codes), we could further assume that citations by the patent office might be retrieved by the IPC filtering. This hypothesis became a motivation for NTCIR-6 to divide up the relevant documents according to the number of shared IPC codes with a search topic (Fujii et al. 2007b).

Ricoh used the IPC codes for both filtering and pseudo-relevance feedback (Itoh 2004, 2005). In the latter usage, they first retrieved documents and extracted IPC codes from the top-ranked documents; then, they filtered out the retrieved documents which did not share any of the extracted IPC codes.

4.3.3 Patent Classification Task: NTCIR-5, NTCIR-6

4.3.3.1 Data Collections

Patent classification tasks were performed in NTCIR-5 (Iwayama et al. 2005) and NTCIR-6 (Fujii et al. 2007b). Table 4.2 summarizes the test collections of the NTCIR patent classification tasks.

The training documents in NTCIR-5 and NTCIR-6 consisted of Japanese (unexamined) patent applications published during 1993–1997 and their English abstracts. Themes and F-terms for these documents were also released. As for test documents in NTCIR-5, 2,008 documents were released for the theme classification task, while 500 were released for the F-term classification task. The documents were selected from Japanese (unexamined) patent applications published in 1998 and 1999. Five themes were selected in the F-term classification task.

In NTCIR-6, only the F-term classification task was performed. We increased the number of themes to 108, and the test documents to 21,606.

Table 4.2 NTCIR test collections for patent classification

Task		NTCIR-5		NTCIR-6
		Theme	F-term	F-term
Test documents		2,008	2,562 (5 themes)	21,606 (108 themes)
Training documents	Patent applications	1993–2002 ja		
	Abstracts	1993–2002 en		

4.3.3.2 Submissions

Each participant in NTCIR-5 submitted a ranked list of themes (at maximum 100) for each test document in the theme classification task and a ranked list of F-terms (at maximum 200) for each test document in the F-term classification task. Note that the participants were given the themes of each test document in the F-term classification task.

4.3.3.3 Evaluation

MAP and F-measure were used in the evaluation. To calculate the F-measure, participants were requested to submit a confident set of themes or F-terms for each test document.

4.3.3.4 Participants

Four groups submitted results to the theme classification task in NTCIR-4. Theme classification is similar to classifying patent applications into IPC sub-classes; k-Nearest Neighbor (k-NN) and naïve Bayes classifiers were popular methods, and the participants used these methods in the task (Kim et al. 2005; Tashiro et al. 2005).

Three groups participated in the F-term classification task in NTCIR-5 and six in NTCIR-6. Some groups used support vector machine (SVM) (Tashiro et al. 2005; Li et al. 2007) in addition to k-NN (Murata et al. 2005) and naïve Bayes (Fujino and Isozaki 2007) classifiers. The results suggested that feature selection had a greater influence on classification effectiveness than the choice of classifier. Since patent applications have several components including abstract, claim, technological field, purpose, embodiments, etc., we have many options for which components should be used as the source of features.

4.3.4 Patent Mining Task: NTCIR-7, NTCIR-8

Patent mining tasks were performed in NTCIR-7 (Nanba et al. 2008) and NTCIR-8 (Nanba et al. 2010). Table 4.3 summarizes the test collections of the patent mining tasks.

The purpose of the patent mining task was to create technical trend maps from a set of research papers and patents. Table 4.4 shows an example of a technical trend map. In this map, research papers and patents are classified in terms of elemental technologies and their effects.

Table 4.3 NTCIR test collections for patent mining

Task		NTCIR-7	NTCIR-8	
		Classification	Classification	Map creation
Test document		879	549	200
Training documents	Patent applications	1993–2002 ja, en		
	Abstracts of research papers	1988–1999 ja, en		

Table 4.4 Example of a technical trend map created from a set of research papers and patents

	Effect 1	Effect 2	Effect 3
Technology 1	[AA 1993] [US Pat. XX/XXX]		[BB 2002]
Technology 2	[CC 2000]		
Technology 3		[US Pat. YY/YYYY]	[US Pat. ZZ/ZZZZ] [JP Pat. WW/WWWW]

Two steps were used to create a technical trend map:

- (Step 1) For a given field, collect research papers and patents written in various languages.
- (Step 2) Extract elemental technologies and their effects from the documents collected in Step 1 and classify the documents in terms of the elemental technologies and their effects. Example of elemental technologies and their effects will be shown in Sect. 3.4.4.

Two subtasks were conducted in each step:

- Classify research paper abstracts.
- Create a technical trend map.

We describe the details of these subtasks below.

4.3.4.1 Research Paper Classification Subtask

The goal of this subtask was to classify research paper abstracts in accordance with the IPC system, which is a standard hierarchical patent classification system used around the world. One or more IPC codes are manually assigned to each patent, aiming for effective patent retrieval.

Task

This task involved assigning one or more IPC codes at the subclass, main group, and subgroup levels to a given topic, expressed in terms of the title and abstract of a research paper.

The following tasks were conducted.

- Japanese: classification of Japanese research papers using patent data written in Japanese.
- English: classification of English research papers using patent data written in English.

Data Collection

We created English and Japanese topics (titles and abstracts) and their correct classifications (IPC codes extracted from patents). On average, 1.6, 1.9, and 2.4 IPC codes were assigned at the subclass, main group, and subgroup levels, respectively, to each topic. In NTCIR-7, we randomly assigned 97 topics to the dry run and the remaining 879 topics to the formal run. In NTCIR-8, we assigned 95 topics to the dry run and the remaining 549 topics to the formal run. The dry run data were provided to the participating teams as training data for the formal run. Patents with IPC codes were also provided as additional training data.

Submission

Participating teams were asked to submit one or more runs, each of which contained ranked lists of IPC codes for each topic.

Evaluation

MAP, recall, and precision were used in the evaluation.

Participants

In NTCIR-7, we had 24 participating systems for the Japanese subtask, 20 for the English subtask, and five for the cross-lingual subtask. As far as the number of groups is concerned, we had 12 participating groups from universities and companies. In NTCIR-8, there were 71 participating systems for the Japanese subtask, 24 for the English subtask, and nine for the cross-lingual subtask. There were six participating groups.

Most participating teams employed the k-Nearest Neighbor (k-NN) method, which is a comparatively easy way of dealing with a large number of categories, because the classification is based only on extracting similar examples, with no training process being required. Furthermore, the k-NN method is itself a ranking, which enables it to be applied directly to the IPC code ranking. In NTCIR-7, Xiao et al. (2008) used the k-NN framework and various similarity calculation methods and re-ranking methods were examined.

4.3.4.2 Technical Trend Map Creation Subtask

Task

This task was conducted in NTCIR-8. The goal of this subtask was to extract expressions of elemental technologies and their effects from research papers and patents. We defined the tag set for this subtask as follows:

- TECHNOLOGY including algorithms, tools, materials, and data used in each study or invention.
- EFFECT including pairs of ATTRIBUTE and VALUE tags.
- ATTRIBUTE and VALUE including effects of a technology that can be expressed by a pair comprising an attribute and a value.

For example, suppose that the sentence “Through closed-loop feedback control, the system could minimize the power loss.” is given to a system. In this case, the system was expected to output the following tagged sentence: “Through <TECHNOLOGY>closed-loop feedback control</TECHNOLOGY>, the system could <EFFECT><VALUE>minimize</VALUE> the <ATTRIBUTE>power loss</ATTRIBUTE> </EFFECT>.”

The following tasks were conducted:

- Japanese: extraction of technologies and their effects from research papers and patents written in Japanese.
- English: extraction of technologies and their effects from research papers and patents written in English.

Data Collection

Sets of topics with manually assigned TECHNOLOGY, EFFECT, ATTRIBUTE, and VALUE tags were used for the training and evaluation. Here, we asked a human subject to assign these tags to the following four text types:

- Japanese research papers (500 abstracts)
- Japanese patents (500 abstracts)
- English research papers (500 abstracts)
- English patents (500 abstracts)

Then, for each text type, We randomly selected 50 texts for the dry run and 200 texts for the formal run. We provided the remaining 250 texts to the participating teams as training data.

Submission

The teams were asked to submit texts with automatically annotated tags.

Evaluation

Recall, precision, and F-measure were used in the evaluation.

Participants

In NTCIR-8, there were 27 participating systems for the Japanese subtask and 13 for the English subtask. There were nine participating teams of universities and companies. For example, Nishiyama et al. (2010) used a system that applied a domain-adaptation method on both research papers and patents and confirmed its effectiveness.

4.4 Contributions

This section chronologically summarizes NTCIR’s contributions to research activities on patent information retrieval. Figure 4.1 shows an overview.

4.4.1 Preliminary Workshop

In 2000, the Workshop on Patent Retrieval was co-located with the ACM SIGIR Conference on Research and Development in Information Retrieval (Kando and Leong 2000). This was the first opportunity for researchers and practitioners associated with patent retrieval to exchange knowledge and experience. The outcome of this workshop motivated researchers to foster research and development in patent retrieval by developing large test collections.

4.4.2 Technology Survey

Following the workshop in 2000, NTCIR-3 was organized as the first evaluation workshop focusing on patent information retrieval (2001–2002). The task was a technology survey.

Since patent offices publish patent applications in public, information retrieval, and natural language processing researchers can use them as a resource. The test collection constructed for NTCIR-3 was unique in that it contained not only patent applications but also search topics and their relevant documents; these were created

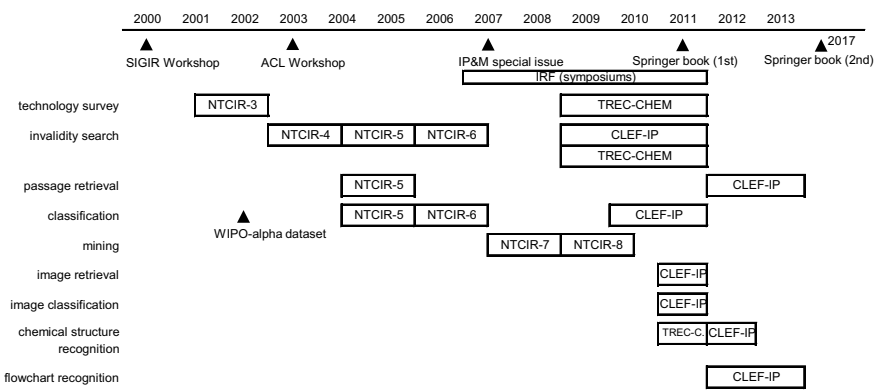


Fig. 4.1 History of research activities on patent information processing

and assessed by human experts. It was the first test collection for patent information retrieval with a large number of documents.

Here we should note that other workshops included technology survey tasks similar to the one performed in NTCIR-3, including the TREC-CHEM tracks in 2009 (Lupu et al. 2009), 2010 (Lupu et al. 2010) and 2011 (Lupu et al. 2011a). These tasks focused on research and development in the chemical domain, in which patent information plays important role.

4.4.3 Collaboration with Patent Experts

The organizers of NTCIR-3 and NTCIR-4 collaborated with patent experts, who were JIPA members, in constructing the test collections. The JIPA members created the search topics and they also collected and assessed relevant documents. The organizers and the JIPA members met once a month to discuss the task design. The participants and the JIPA members also shared knowledge and experiences at round-table meetings and tutorials. These activities helped to build bridges between information retrieval researchers and patent searchers.

4.4.4 Invalidity Search

NTCIR-4 (2003–2004) was the first workshop to include an invalidity search. Invalidity search is truly patent-specific work, and the organizers carefully designed the task with JIPA members.

In NTCIR-4, we examined the issue of whether it was possible to use only citations as relevant documents when evaluating submitted runs. While the NTCIR-4 collection included an exhaustive collection of relevant documents, the NTCIR-5 and NTCIR-6 collections had only citations. Moreover, since the topics and the documents were the same in the NTCIR4, NTCIR-5, and NTCIR-6 collections, researchers can compare their retrieval methods under the different ways of identifying relevant documents.

Invalidity search tasks were continuously organized in CLEF-IP in 2009 (Roda et al. 2009), 2010 (Piroi and Tait 2010) and 2011 (Piroi et al. 2011), and TREC-CHEM in 2009 (Lupu et al. 2009), 2010 (Lupu et al. 2010) and 2011 (Lupu et al. 2011a), under the name “prior art search task”.

The passage retrieval task in NTCIR-5 (2004–2005) was the first attempt to evaluate handling of passages in invalidity search. A passage retrieval task was revisited in CLEF-IP in 2012 (Piroi et al. 2012) and 2013 (Piroi et al. 2013) in a more challenging setting. In NTCIR, a relevant document to a search topic was given and the purpose was to find relevant passages in the given relevant document. On the other hand, in CLEF-IP, relevant passages were directly retrieved based on the claims in the search topic.

4.4.5 *Patent Classification*

The WIPO-alpha collection, released in 2002, was the first test collection for patent classification. It consisted of 75,250 English patent documents labeled with IPC codes. Many research papers on patent classification have used WIPO-alpha (Fall et al. 2003).

The NTCIR collections released by the classification tasks (2004–2007) were not for IPC, but for the classification codes used in the JPO, i.e., F-terms. F-terms re-classify a specific technical field of IPC from a variety of viewpoints, such as purpose, means, function, and effect.

The CLEF-IP classification tasks in 2010 (Piroi and Tait 2010) and 2011 (Piroi et al. 2011) released test collections on IPC codes; these were larger than the WIPO-alpha collection. The released documents totaled 2.6 million in 2010 and 3.5 million in 2011.

4.4.6 *Mining*

Patent mining tasks were performed in NTCIR-7 and NTCIR-8, and similar tasks were conducted in subsequent research (Gupta and Manning 2011; Tateisi et al. 2016). Gupta and Manning (2011) proposed a method to assign FOCUS (an article's main contribution), DOMAIN (an article's application domain), and TECHNIQUE (a method or a tool used in an article) tags to abstracts in the ACL Anthology² for the purpose of identifying technical trends. Tateisi et al. (2016) constructed a corpus for analyzing the semantic structures of research articles in the computer science domain.

Since February 2019, JDream III³ has provided a new service for retrieving research papers using IPC codes. This service assigns IPC codes at the main group level to each research paper by using Nanba's method (Nanba 2008), which is based on the k-NN method.

4.4.7 *Workshops and Publications*

The organizers of the NTCIR tasks organized the ACL Workshop on Patent Corpus Processing in 2003 and edited a special issue on patent processing in Information Processing & Management in 2007 (Fujii et al. 2007a).

The Information Retrieval Facility (IRF), which was a not-for-profit research institution based in Vienna, Austria, organized a series of symposia between 2007 and 2011 to explore reasons for the knowledge gap between information retrieval

²<https://www.aclweb.org/anthology/>.

³<https://jdream3.com/>.

researchers and patent search specialists. The symposia were followed by publication of two editions of a book in 2011 (Lupu et al. 2011b) and 2017 (Lupu et al. 2017) introducing studies by information retrieval researchers and patent experts.

These activities contributed to the research trends in the communities of information retrieval and natural language processing.

4.4.8 CLEF-IP and TREC-CHEM

The NTCIR project ended with NTCIR-8 (2009–2010), and it left behind several unaddressed issues. Firstly, while NTCIR-3 and NTCIR-4 released multi-lingual search topics in English, Korean and Chinese, as well as English abstracts over the course of ten years and NTCIR-6 included an English retrieval task using patent applications published by USPTO, the workshops focused on Japanese and the multi-lingual resources were not widely used. This meant there were no serious evaluations of multi-lingual or cross-lingual patent retrieval. Secondly, the tasks ignored images, formulas, and chemical structures, despite the fact that these are important pieces of information for judging relevance in some domains.

The above issues that the NTCIR project did not address were investigated in CLEF-IP (2009–2013) and TREC-CHEM (2009–2011). Both were annual evaluation workshops (campaigns) on patent information retrieval. CLEF-IP had tasks for prior art search, passage retrieval, and patent classification. The tasks were similar to the NTCIR tasks, but most resources were from the European Patent Office, covering English, French, and German; hence, the CLEF-IP tasks were inherently multi/cross-lingual. In addition, CLEF-IP performed completely new tasks, including ones on image-based retrieval (Piroi et al. 2011), image classification (Piroi et al. 2011), flowchart/structure recognition (Piroi et al. 2012, 2013), and chemical structure recognition (Piroi et al. 2012). TREC-CHEM also had tasks for prior art search and technology survey. The TREC-CHEM tasks were challenging and focused on the chemical domain, which has many formulae and images. The image-to-structure task (Lupu et al. 2011a) in TREC-CHEM was the first one to include chemical structure recognition. TREC-CHEM used resources from USPTO.

References

- Fall C, Torcsvari A, Benzineb K, Karetka G (2003) Automated categorization in the international patent classification. *ACM SIGIR Forum* 37(1):10–25
- Fujii A (2007) Integrating content and citation information for the NTCIR-6 patent retrieval task. In: *Proceedings of the 6th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access*
- Fujii A, Ishikawa T (2004) Document structure analysis in associative patent retrieval. In: *Proceedings of the 4th NTCIR workshop on research in information access technologies information retrieval, question answering and summarization*

- Fujii A, Ishikawa T (2005) Document structure analysis for the NTCIR-5 patent retrieval task. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Fujii A, Iwayama M, Kando N (2004) Overview of patent retrieval task at NTCIR-4. In: Proceedings of the 5th NTCIR workshop on research in information access technologies information retrieval, question answering and summarization
- Fujii A, Iwayama M, Kando N (2005) Overview of patent retrieval task at NTCIR-5. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Fujii A, Iwayama M, Kando N (2007a) Introduction to the special issue on patent processing. *Inf Process Manag* 45(3):1149–1153
- Fujii A, Iwayama M, Kando N (2007b) Overview of the patent retrieval task at the NTCIR-6 workshop. In: Proceedings of the 6th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Fujino A, Isozaki H (2007) Multi-label patent classification at NTT communication science laboratories. In: Proceedings of the 6th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Gupta S, Manning GD (2011) Analyzing the dynamics of research by extracting key aspects of scientific papers. In: Proceedings of the 5th international joint conference on natural language processing
- Itoh H (2004) NTCIR-4 patent retrieval experiments at RICOH. In: Proceedings of the 4th NTCIR workshop on research in information access technologies information retrieval, question answering and summarization
- Itoh H (2005) NTCIR-5 patent retrieval experiments at RICOH. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Itoh H, Mano H, Ogawa Y (2003) Term distillation for cross-DB retrieval. In: Proceedings of the 3rd NTCIR workshop on research in information retrieval, automatic text summarization and question answering
- Iwayama M, Fujii A, Fujii A, Takano A (2003) Overview of patent retrieval task at NTCIR-3. In: Proceedings of the 3rd NTCIR workshop on research in information retrieval, automatic text summarization and question answering
- Iwayama M, Fujii A, Kando N (2005) Overview of classification subtask at NTCIR-5 patent retrieval task. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Kando N, Leong M (2000) Workshop on patent retrieval SIGIR 2000 workshop report. *ACM SIGIR Forum* 34(1):28–30
- Kim JH, Huang JX, Jung HY, Choi KS (2005) Patent document retrieval and classification at KAIST. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Konishi K (2005) Query terms extraction from patent document for invalidity search. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Konishi K, Kitauchi A, Takaki T (2004) Invalidity patent search system of NTT DATA. In: Proceedings of the 5th NTCIR workshop on research in information access technologies information retrieval, question answering and summarization
- Li Y, Bontcheva K, Cunningham H (2007) SVM based learning system for F-term patent classification. In: Proceedings of the 6th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Lupu M, Piroi F, Huang X, Zhu J, Tait J (2009) Overview of the TREC 2009 chemical IR track. In: Proceedings of TREC 2009

- Lupu M, Tait J, Huang J, Zhu J (2010) TREC-CHEM 2010: Notebook report. In: Proceedings of TREC 2010
- Lupu M, Gurulingappa H, Filippov I, Jiashu Z, Fluck J, Zimmermann M, Huang J, Tait J (2011a) Overview of the TREC 2011 chemical IR track. In: Proceedings of TREC 2011
- Lupu M, Mayer K, Tait J, Trippe AJ (eds) (2011b) Current challenges in patent information retrieval. Springer, Berlin
- Lupu M, Mayer K, Kando N, Trippe AJ (eds) (2017) Current challenges in patent information retrieval, 2nd edn. Springer, Berlin
- Mase H, Iwayama M (2007) NTCIR-6 patent retrieval experiments at hitachi. In: Proceedings of the 6th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Mase H, Matsubayashi T, Ogawa Y, Iwayama M, Oshio T (2004) Two-stage patent retrieval method considering claim structure. In: Proceedings of the 4th NTCIR workshop on research in information access technologies information retrieval, question answering and summarization
- Mase H, Matsubayashi T, Ogawa Y, Yayoi T, Sato Y, Iwayama M (2005) NTCIR-5 patent retrieval experiments at hitachi. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Murata M, Kanamaru T, Shirado T, Isahara H (2005) Using the k nearest neighbor method and BM25 in the patent document categorization subtask at NTCIR-5. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Nanba H (2008) Hiroshima City University at NTCIR-7 patent mining task. In: Proceedings of the 7th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Nanba H, Fujii A, Iwayama M, Hashimoto T (2008) Overview of the patent mining task at the NTCIR-7 workshop. In: Proceedings of the 7th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Nanba H, Fujii A, Iwayama M, Hashimoto T (2010) Overview of the patent mining task at the NTCIR-8 workshop. In: Proceedings of the 8th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Nishiyama R, Tsuboi Y, Unno Y, Takeuchi H (2010) Feature-rich information extraction for the technical trend map creation. In: Proceedings of the 8th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access
- Piroi F, Tait J (2010) CLEF-IP 2010: retrieval experiments in the intellectual property domain. In: Working notes for CLEF-2010 conference
- Piroi F, Lupu M, Hanbury A, Zenz V (2011) CLEF-IP 2011: retrieval experiments in the intellectual property domain. In: Working notes for CLEF-2011 conference
- Piroi F, Lupu M, Hanbury A, Sexton A, Magdy W, Filippov I (2012) CLEF-IP 2012: retrieval experiments in the intellectual property domain. In: Working notes for CLEF-2012 conference
- Piroi F, Lupu M, Hanbury A (2013) Overview of CLEF-IP 2013 lab. In: CLEF 2013 Proceedings of the 4th international conference on information access evaluation. Multilinguality, multimodality, and visualization
- Roda G, Tait J, Piroi F, Zenz V (2009) CLEF-IP 2009: Retrieval experiments in the intellectual property domain. In: Multilingual information access evaluation I. Text retrieval experiments 10th workshop of the cross-language evaluation forum, CLEF 2009
- Tashiro T, Rikitoku M, Nakagawa T (2005) Justsystem at NTCIR-5 patent classification. In: Proceedings of the 5th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access

- Tateisi Y, Ohta T, Miyao Y, Pyysalo S, Aizawa A (2016) Typed entity and relation annotation on computer science papers. In: Proceedings of the 10th international conference on language resources and evaluation (LREC 2016)
- Xiao T, Cao F, Li T, Song G, Zhou K, Zhu J, Wang H (2008) Knn and re-ranking models for english patent mining at NTCIR-7. In: Proceedings of the 7th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

