

NTCIR タスク参加者用テストコレクション利用許諾に関する覚書

(NTCIR ワークショップ参加者用)

大学共同利用機関法人 情報・システム研究機構 国立情報学研究所（以下「甲」という）と
（以下「乙」という）は、NTCIR ワークショップにおいて甲が提供する
「NTCIR-8 高度言語横断情報アクセス タスク参加者用テストコレクション」に関して、以下の通りの覚書
を結ぶこととする。

第一条（データの内容）

1. 「文書データ」とは、別紙細則 1 に定めるデータのことである。
2. 「タスクデータ」とは、別紙細則 2 に定めるデータのことである。
3. 「未加工提出結果データ」とは、甲が主催する NTCIR ワークショップにおいて、甲が設定した課題について、当該タスクの参加者が提出したシステムの実行結果の未加工のデータのことである。
4. 「評価データ」とは、「未加工提出結果データ」に対して、あらかじめ定めた基本評価算出ツールを用いて算出して得られたデータのことである。
5. 「テストコレクション」とは、「文書データ」、「タスクデータ」、「未加工提出結果データ」、「評価データ」の総称である。

第二条（権利の帰属）

1. 「文書データ」に関する著作権については、別紙細則 3 に定める。
2. 「タスクデータ」に関する著作権については、別紙細則 4 に定める。
3. 乙が、「テストコレクション」を利用して開発した技術、システムなどに関して生じた知的財産権は、乙に帰属する。
4. 乙から提出されたデータに基づいて、甲が行った分析結果、「テストコレクション」の改良などに関して生じた知的財産権は、甲に帰属する。

第三条（利用許諾）

甲は乙に対して「テストコレクション」の利用を許諾する。

第四条（利用許諾の範囲）

1. 乙は、NTCIR-8 ワークショップが終了する平成 23 年 3 月末日まで、「テストコレクション」を NTCIR ワークショップの課題遂行および課題に関連する研究目的にのみ利用できるものとする。
2. 乙は、NTCIR-8 ワークショップ終了後は「テストコレクション」を情報検索、自然言語処理等に関する研究目的にのみ利用できるものとする。
3. 乙は、「テストコレクション」およびその全体または一部を複製したもの、あるいは、それを復元できる状態に加工されたデータを第三者に対して、売買、貸与、刊行、配布、送信可能化をしてはならない。

第五条（提供の方法）

甲は、別紙細則 5 に定める手段により「テストコレクション」を乙に提供する。

第六条（利用者の範囲）

1. 「テストコレクション」の利用者の範囲は、乙本人および乙と直接共同して研究するグループの構成員に限定されるものとする。
2. 乙は、利用者の名簿を管理し、甲から求めがあった場合は、遅滞なく、これを甲に提出するものとする。

第七条（知見の発表）

1. 乙は、本覚書に違反しない範囲において、「テストコレクション」を利用して得られた知見に関する研究発表を行うことができる。
2. 乙は、研究発表において、自己の研究を記述するために必要な場合に限り、「テストコレクション」に含まれるデータの一部を引用することができる。その際、引用する部分の著作権および出版者等の権利を侵害してはならない。
3. 乙は、「未加工提出結果データ」もしくは「評価データ」を利用した研究成果を発表する場合は、発表論文において、これらのデータの性質と収集法を明確に説明する。
4. 乙は、発表論文に、テストコレクション名及び利用した文書データを明記し、かつ、NTCIR ワークショップの会議論文集と関連する文献を引用するものとする。
5. 乙は、発表論文の書誌事項（掲載資料名、巻号ページ、出版者、発表年月日等）とともに発表論文の別刷りまたはコピーを一部、論文発表の都度、甲に提出するものとする。
6. 乙は、「テストコレクション」を利用して得られたデータの公開については、事前に甲から書面によ

る承認を得ることとする。

7. 乙は、「テストコレクション」を用いた評価結果を商品の広告、宣伝などの営利目的、および誹謗・中傷に用いてはならない。

第八条（覚書の有効期間）

1. 本覚書の有効期間は、覚書締結日より平成 23 年 3 月 31 日までとする。ただし、有効期間満了日の 1 ヶ月前までに、甲または乙から本覚書を更新しない旨の書面による申し出がない場合には、一年間継続するものとし、以後も同様とする。
2. 前項にかかわらず、一年間継続する乙が「未加工提出結果データ」および「評価データ」の利用を希望する場合は、有効期間満了日の一か月前までに甲に書面により申し出た上、別途覚書を取り交わすこととする。
3. 有効期間または更新期間をさらに更新しない場合は、乙は期間満了後、乙はすべての計算機およびメディアから「テストコレクション」を速やかに消去しなければならない。なお、乙の属する組織または乙の所属に変更の生じた場合は、遅滞なくこれを甲に報告し、必要があれば覚書の取り交わしを改めて行うものとする。

第九条（報告書の提出）

1. 乙は、NTCIR ワークショップの定められた手順に従って、成果報告を甲に提出するものとする。
2. 乙は、有効期間満了日の一ヶ月前までに「テストコレクション」を利用した当該年度の研究活動に関する報告書を甲へ提出するものとする。

第十条（データの利用中止）

1. 乙は、本覚書に違反する利用が行われた場合、甲の申し入れにより、直ちに「テストコレクション」の利用を中止し、すべての計算機およびメディアから、「テストコレクション」およびこれを加工して得られたデータの全てを速やかに消去し、消去した旨の書面を甲に提出しなければならない。
2. 「テストコレクション」の著作権者または利用許諾権者から、個々の文書データまたはタスクデータの利用中止の要請があった場合、乙は、甲の申し入れにより、すべての計算機およびメディアから該当するデータを速やかに消去し、消去した旨の書面を甲に提出しなければならない。

第十一条（免責事項）

甲および「テストコレクション」の著作権者または利用許諾権者は、理由の如何を問わず、乙が「テストコレクション」を利用したことにより生じた不利益について、一切の責任を負わないものとする。

第十二条（協議事項）

本覚書に定めのない事項が生じた場合は、甲乙は誠意を持って協議し、問題を解決するものとする。

第十三条（管轄裁判所）

本覚書に関する一切の紛争については、東京地方裁判所を第一審の専属的合意管轄裁判所とする。

以上、本覚書の成立の証として本書を 2 通作成し、甲乙記名押印の上、各 1 通を保有する。

平成 年 月 日

(甲) 東京都千代田区一ツ橋 2-1-2
大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
NTCIR プロジェクト プロジェクトリーダー
教授 神門典子

(乙) 住所 _____
名称 _____
役職名 _____
氏名 _____

別紙—NTCIR-8 高度言語横断情報アクセス タスク参加者用テストコレクション

(NTCIR ワークショップ 8 参加者用)

細則 1 「文書データ」とは、正式名称「NTCIR-8 高度言語横断情報アクセスタスク 文書データ」と称し、甲が、その著作権者または利用許諾権者から NTCIR ワークショップ 8 における利用のために一定期間参加者に利用させることを許諾されたデータのことであり、表 A に掲げた中から乙によって選択され、構成されたデータのことである。

表 A (必要なものにチェックすること)

- A.1 中国語データセット (繁体字)
UDN2002-2005: United Daily News, United Express, Ming Hseng News, Economic Daily News, Star News (Copyright holder: UDN.COM)
- A.2 日本語データセット
毎日新聞全文記事データベース 2002-2005 (著作権者: 毎日新聞社)
- A.3 システム訓練用中国語データセット (繁体字)
CIRB011 1998-1999: China Times, Commercial Times, China Times Express (Copyright holder: China Times Inc.); Central Daily News (Copyright: Central Daily News); China Daily News (Copyright holder: China Daily News)
CIRB020 1998-1999: United Daily News, Economic Daily News, Min Sheng Daily, United Evening News, Star News (Copyright holder: UDN.COM)
CIRB040 2000-2001: United Daily News, United Express, Ming Hseng News, Economic Daily News (Copyright holder: UDN.COM)
- A.4 システム訓練用 中国語データセット (簡体字)
Lianhe Zaobao 1998-2001 (Copyright holder: Singapore Press Holdings Limited)
- A.5 システム訓練用 日本語データセット
毎日新聞全文記事データベース 1998-2001 (著作権者: 毎日新聞社)

なお、「毎日新聞全文記事データベース 1998-2001」、「毎日新聞記事データベース 2002-2005」および「CIRB011」の利用許諾に関しては、NTCIR ワークショップの定められた手順に従って所定の期日までに、成果報告を甲に提出した参加者に限り、一年間継続するものとし以後も同様とする。ただし、「毎日新聞全文記事データベース 1998-2001」および「毎日新聞記事データベース 2002-2005」の利用期間の更新は最大 5 回を限度とし、利用期間の最終期限は平成 28 年 3 月 31 日までとする。

細則 2

1. 「タスクデータ」とは、正式名称「NTCIR-8 高度言語横断情報アクセスタスク タスクデータ」と称し、「情報検索データ」と「質問応答データ」の総称であり、表 B に掲げた中から乙によって選択され、構成されたデータのことである。

表 B (必要なものにチェックすること)

- NTCIR-5 多言語質問応答タスク 質問応答データ
- NTCIR-6 多言語質問応答タスク 質問応答データ
- NTCIR-3 質問応答タスク 質問応答データ
- NTCIR-4 質問応答タスク 質問応答データ
- NTCIR-5 質問応答タスク 質問応答データ
- NTCIR-6 質問応答タスク 質問応答データ
- NTCIR-3 言語横断検索タスク 情報検索データ
- NTCIR-4 言語横断検索タスク 情報検索データ
- NTCIR-5 言語横断検索タスク 情報検索データ
- NTCIR-6 言語横断検索タスク 情報検索データ
- NTCIR-7 高度な言語横断情報検索と質問応答タスク 質問応答データ・情報検索データ
- NTCIR-8 高度言語横断情報アクセスタスク 質問応答データ・情報検索データ

2. 「情報検索データ」とは、「文書データ」の内容に対して行う検索質問、各検索質問に適合する記事の ID のリスト、および、適合と判断した根拠を示すために文書から抽出した語句の総称である。
3. 「質問応答データ」とは、「文書データ」に含まれる記事データに対して行う質問文、質問文の答えとなる記述、および、答えと判断するために必要な記事の ID とその重要箇所からなるものの総称である。

細則 3 「文書データ」に関する著作権は、表 A に示した各データの権利者に帰属する。

細則 4 「タスクデータ」に関する著作権は、甲に帰属する。「質問応答データ」は、「文書データ」に含まれる記事の内容に関するものであるが、甲が NTCIR ワークショップのために作成したものであり、その内容に関しては「文書データ」の著作権者が責任を負うものではない。

細則 5

1. 甲は、乙に対する「文書データ」の提供を技術的に妥当な手段により行う。
2. 甲は、乙に対する「タスクデータ」の提供をファイル転送などの電子的手段により行う