

## 仕 様 書

### 1. 委託業務題目

2022 年度産業の研究開発に関する基盤的なデータ整備

### 2. 委託業務の目的及び概要

科学技術・学術政策研究所（以下、科政研と略す）では、文部科学省が実施している「政策のための科学」推進事業におけるデータ・情報基盤整備の一環として、客観的データに基づいた科学技術イノベーション政策の形成に資することを目的に、国内営利企業（以下、企業と略す）の知的財産、研究開発、事業等に関するデータを体系的かつ有機的に結合し、産業セクターにおける研究開発やイノベーションの測定を可能とするデータ・情報基盤の整備を行っている。

本委託業務は、これまでに整備した産業の研究開発に関するデータ・情報基盤を継続的・発展的に利活用可能な基盤とすることを目的に実施する。

委託する業務は、データ・情報基盤の中核となる NISTEP 企業名辞書（以降、企業名辞書と略す）の掲載情報及び関連接続テーブルの最新化を行なう「定常的实施事項」、並びに産業セクターの科学技術イノベーション分析等により有効に利活用できるデータ・情報基盤とするための「アドホック的实施事項」の二つを実施する。

### 3. 委託業務の内容

#### 3. 1 定常的实施事項

##### 3. 1. 1 企業名辞書の最新化

企業名辞書は、企業に関する、法人番号、商号変更・合併等の沿革、所在地、緯度経度、規模、業種、ベンチャーなど多岐に渡る企業情報テーブルで構成するリレーショナルデータベース（以降、RDB と略す）であり、特許データ、財務データ、各種企業調査データなど、研究者が分析に必要とする任意の情報と接続して利用することができる。

3. 1. 1 では、次の（１）、（２）の項目について状況確認を行い、変化がある場合、企業名辞書テーブルに当該事象に関するデータの追加を行う。尚、変化前のデータはパネル化し保持することが原則であり、上書き更新せずに最新データを更新日とともに新規に追加する。また、（１）、（２）の状況確認は業務受託者の本作業時点の企業情報で行う。

##### （１）企業状況の確認とデータ更新

科政研より受託者に提供する企業名辞書<sup>1</sup>（本委託業務のベースとする企業名辞書。以降「企業名辞書 [Base]」と称す）に掲載する企業に関して、以下の①～⑤の状況の有無を確認し、有の場合、当該企業情報の更新又は追加を行う。無の場合、確認日を更新し、確認時点で企業状況に変化がなかったことを明示する。

- ①商号に変更がある場合
- ②企業の統合・再編等が行われた場合
- ③倒産、清算等により会社解散が行われた場合
- ④企業の連結関係に変化が生じた場合

---

<sup>1</sup> 本仕様書記述時点で検証中であり未公開版である 2022\_1 版

⑤資本金、従業員数、主業、所在地、企業 URL、その他の企業情報に変更がある場合

## (2) 企業名辞書掲載条件新規クリア企業の追加

次の①、②のいずれかの条件に該当する企業が企業名辞書 [Base] に未掲載であれば、当該企業を新規に追加する。当該企業が商号変更、合併などの沿革がある場合は、それら沿革企業の情報も合わせて追加する。尚、①及び②の条件チェックにおいて、特許出願情報の利用は IIP パテントデータベース(受託者の作業時点で公開されている版)<sup>2</sup>を前提に実施すること。

### ① 特許出願累積数 100 件以上の企業

1970 年以降に特許出願した累積数が 100 件以上ある企業である。ここで、累積数は企業の沿革を考慮して算出した数である。例えば、A という商号の企業が B に変更したと仮定して、A 社の時代に 80 件、B 社となってから 20 件超出願したとするならば、B 社としての累積数は 100 件超となり掲載条件を満たしたことになる。

### ② 出願件数伸び率が設定閾値を超えた企業

出願件数の伸び率は、1970 年以降の 3 年、5 年、7 年の各期間で 1 年ごと移動させた線形フィット値で評価する。ここで、商号変更、合併などがある企業は、沿革を考慮した出願数を用いて伸び率を算出する。

線形フィットの閾値は、それぞれ 8、21/5、75/28 として、閾値を超える伸び率を持つ未掲載企業があれば新規に追加する。

## 3. 1. 2 東証市場区分見直しに伴うデータ更新

上場企業の市場区分は、2022 年 4 月 4 日付けにて市場第一部、市場第二部等の区分からプライム市場等の新市場区分に移行した。ここでは、企業名辞書 [Base] について、上場企業の新市場区分への移行等に伴う市場区分データの更新を行う。また、企業名辞書 [Base] に未掲載の新規上場企業があれば追加する。

市場区分データの更新は先行的に業務を実施し、更新した企業名辞書は他の成果物とは別に 2022 年 9 月末日を目標に提出を行うこと。

## 3. 1. 3 外部データベースとの接続テーブルの最新化

企業名辞書と外部データベースとの接続性を向上させ統合的データ利用を可能にする目的で実施する。

### (1) IIP パテントデータベースとの接続テーブル

本委託業務で作成した企業名辞書(2023\_1 版と仮称)<sup>3</sup>と IIP パテントデータベース(作業時点公開版)とを繋ぐ接続テーブルを作成する。

企業名辞書と IIP パテントデータベースに登場する企業の同一性の判定は、企業名称(法人格を削除した企業名称文字列)、法人格コード(前株・後株などを表すコード)、住所コード(市区町村レベルの 5 桁コード)の論理積、および、特許庁より付与された出願人番号を考慮して行う。但し、別途、科政研より提供する接続テーブルの接続情報は維持し、新たな接続情報を追加する形態で作成する。

<sup>2</sup> (一財)知的財産研究所が特許統計の研究普及の目的で公開する特許データベース。本仕様書記述時点の版は 2020 年公開版。

<sup>3</sup> 本仕様書の 3. 1～3. 2. 2 の作業により企業名辞書 [Base] に企業情報の追加・改訂を行った版

接続テーブルは、comp\_id(企業番号)、ida\_seq(ida(出願番号)と seq(出願人記載順序)をアンダーバーで接続)及び adate (出願日) の3つのフィールドで構成する。

## (2) NISTEP 大学・公的機関名辞書との接続テーブルの最新化

NISTEP 大学・公的機関名辞書<sup>4</sup>に収録された企業(論文著者の所属企業)を、企業名辞書に収録し、企業名辞書を介して科学論文と特許出願を企業レベルで統合的に接続する。

別途、科政研より提供する NISTEP 大学・公的機関名辞書(最新版)に収録する企業について、企業名辞書 [Base] に未掲載企業について追加を行う。尚、NISTEP 大学・公的機関名辞書には 4500 社が収録されるが、殆どが企業名辞書に掲載済みであり、未掲載企業は 100 企業未満となる見込みである。

## 3. 2 アドホック的实施事項

### 3. 2. 1 企業情報からの所在地情報の自動抽出(実用化に向けたK/Hの蓄積)

企業の所在地情報は、同名異企業が多数存在する中で該当社を選別する際の過誤を防ぐ情報として利用できる。一方、企業の所在地は移転等により変更が頻繁に起こること、また、所在地の抽出元となる企業沿革には不要な情報や主語が明確でないなどの不定形記述であることも珍しくなく、人手による作業でも抽出に手間を要する。

前年度は、この抽出を自動化する可能性を探るべく深層学習自然言語処理プログラム(BERT<sup>5</sup>)を用いて有価証券報告書からの住所の固有表現抽出の試行を行った。今年度は、試行で明らかになった問題点を改善し、実用化乃至は実用化に向けた処理ノウハウを蓄積する期間とする。

尚、本委託の応札者の申し出に応じ、前年度の試行結果に関する報告書を提供する。また、自然言語処理に関して、応札者が手法等に知見を有する場合は提案をお願いする。

### 3. 2. 2 企業名辞書の今後に向けた在り方を議論する有識者会議の運営

産業セクターのイノベーション研究など企業名辞書を活用する有識者により今後の企業名辞書の在り方を討議することを目的とする会議の運営を行う。

会議の開催は3回程度、有識者は4乃至5名で構成するオンライン形式の会議とする。有識者の人選及び承諾手続き等は科政研と協議し行う。

尚、謝金及びその処理業務は科政研で処置するので運營業務の範囲外とする。

その他、会議の運営に係る主要業務は次の通りである。

#### ・会議運營業務

- ①開催日の設定、Zoom 等リモート参加環境の設定
- ②有識者との連絡調整
- ③議事進行
- ④議事録作成

尚、最終会議において、有識者からの提言、助言等が明確化されるよう議事進行に配慮すること。

---

<sup>4</sup> 科政研が公開する研究活動を行っている我が国の機関(約1万3千機関)を掲載したデータベース。大学(大学共同利用機関、短期大学、高等専門学校を含む)及び公的研究機関(国の機関、独立行政法人等[※])を中心として掲載する。[※ 国立研究開発法人を含む。]

<sup>5</sup> BERT は 2018 年に Google から発表された自然言語処理技術であり、コンテンツの文脈を読むことに優れた特徴を持つといわれる。

#### 4. 委託業務成果報告書および処理マニュアルの作成

- ・本委託業務の成果を明示化するために、「3. 委託業務の内容」の成果の目録及び各成果の概要を記述した委託業務成果報告書を作成する。
- ・「3. 委託業務の内容」の成果を得るための処理過程を明示した処理マニュアルを作成する。
- ・「3. 委託業務の内容」の実施にあたっては、科政研と定期的に打ち合わせを行い、業務の進捗報告等を行うこと。また、仕様について不明な点が生じた場合や、データ処理手法等についての判断が必要な場合には、科政研の指示を仰ぐこと。

#### 5. 委託業務実施期間

契約日から 2023 年 3 月 24 日(金)

#### 6. 成果物

本委託業務の実施結果として、以下に記載する成果物を提出すること。尚、成果物は、原則、電子媒体で提出願う。ただし、以下において[電子媒体及び紙媒体]としたものは電子媒体及び紙媒体(各 1 部)を提出すること。

- ① 企業名辞書(2023\_1 版) RDB 及びエクセルファイル [電子媒体]  
但し、東証市場区分を改訂した版は 2022 年 9 月末日を目標に提出を行うこと。
- ② 企業名辞書(2023\_1 版)のフェイスシート<sup>6</sup> [電子媒体]
- ③ IIP パテントデータベース(2020 年公開版又は新版)のフェイスシート [電子媒体]
- ④ フェイスシート作成用住所と住所コード対応表 [電子媒体]
- ⑤ 接続テーブル作成用プログラム
- ⑥ 企業名辞書(2023\_1 版)と IIP パテントデータベース(2020 年公開版又は新版)との接続テーブル [電子媒体]
- ⑦ 自然言語処理プログラムと利用マニュアル [電子媒体]
- ⑧ 有識者会議議事録 [電子媒体]
- ⑨ 委託業務成果報告書 [電子媒体(CD-R)及び紙媒体]
- ⑩ 上記データの処理過程と成果物のデータ構造等を記載した処理マニュアル [電子媒体及び紙媒体]
- ⑪ 上記成果物の作成のために本委託で作成した処理プログラム及びプログラムマニュアル [電子媒体及び紙媒体]

#### 7. 応札者に求める要求要件

##### (1) 要求要件の概要

- ①本委託業務に係る応札者に求める要求要件は、「(2) 要求要件の詳細」による。
- ②要求要件は必須の要求要件と必須以外の要求要件がある。
- ③「\*」の付してある項目は必須の要求要件であり、最低限の要求要件を示しており、技術審査においてこれを満たしていないと判断がなされた場合は不合格として落札決定の対象から除外される。
- ④必須以外の要求要件は、満たしていれば望ましい要求要件であるが、満たしていなくて

---

<sup>6</sup> データベース間で企業名を寄せし、接続テーブルを生成するために使用する企業名称、法人格コード、住所コードからなるデータシートである

も不合格とならない。

- ⑤これらの要求要件を満たしているか否かの判断及びその他提案内容の評価等は、技術審査委員会において行う。尚、総合評価落札方式に係る評価基準は別添の総合評価基準に基づくものとする。

(2) 要求要件の詳細

別添の総合評価基準の「評価項目及び得点配分基準」と同様。

## 8. 無償貸付を行う資料

科政研は、受託者に対し、下記の資料の無償貸付を行う。

- ① 企業名辞書(2022\_1 版) (3 項の企業名辞書 [Base] と同一版)
- ② 企業名辞書(2022\_1 版)と IIP パテントデータベース(2020 年公開版)との接続テーブル
- ③ 企業名辞書(ver. 2022\_1)のフェイスシート
- ④ IIP パテントデータベース(2020 年公開版)のフェイスシート
- ⑤ 接続テーブル作成用プログラム
- ⑥ 企業情報からの所在地情報の自動抽出試行結果の報告書
- ⑦ 2020 年度委託業務成果報告書、同処理マニュアル、同プログラムマニュアル
- ⑧ IIP パテントデータベース(2020 年公開版又は新版)

## 9. 守秘義務

受託者は、本委託業務の実施で知り得た非公開の情報を如何なる者にも漏洩してはならない。

受託者は、本委託業務に係わる情報を他の情報と明確に区別して、善良な管理者の注意をもって管理し、本委託業務以外に使用してはならない。

受託者は、本委託業務終了後速やかに、科政研が貸付した資料を返却するとともに、貸付したデータのバックアップ等を消去すること。

### 10. 届出義務

受託者は、提案書類の提出後、技術審査の日までにおいて、第三者から資格や認定の取消しを受けるなどの後発事象により、提案書の内容に変更が生じることが判明した場合には、速やかに委託者に届け出ること。(女性の職業生活における活躍の推進に関する法律に基づく認定など)

### 11. 賃上げを実施する企業に関する指標に係る留意事項

委託者は、受託者が賃上げを実施する企業に関する指標における加点を受けた場合、受託者の事業年度等(事業年度及び暦年をいう。)が終了した後、表明した率の賃上げを実施したことを以下の手法で確認する。

- ・別添の総合評価基準別紙における【5-1-1】の場合は、賃上げを表明した年度及びその前年度の法人事業概況説明書の「10 主要科目」のうち「労務費」、「役員報酬」及び「従業員給料」の合計額を「4 期末従業員等の状況」のうち「計」で除した金額を比較する。
- ・別添の総合評価基準別紙における【5-1-2】の場合は、給与所得の源泉徴収票等の法定調書合計表の「1 給与所得の源泉徴収票合計表(375)」の「A 俸給、給与、賞与等の総額」の「支払金額」欄を「人員」で除した金額により比較する。

※ 中小企業等にあつては、上記の比較をすべき金額は、【5-1-1】の場合は「合計額」と、【5-1-2】の場合は「支払金額」とする。

加点を受けた受託者は、確認のため必要な書類を速やかに委託者に提出すること。ただし、前述の書類により賃上げ実績が確認できない場合であっても、税理士又は公認会計士等の第三者により上記基準と同等の賃上げ実績を確認することができる書類であると認められる書類等をもって代えることができる。

上記の確認を行った結果、加点を受けた受託者が表明書に記載した賃上げ基準に達していない場合又は本制度の趣旨を意図的に逸脱している場合においては、当該事実判明後の総合評価落札方式において所定の点数（加点の点数を超える点数）を減点するものとする。詳細は表明書裏面の（留意事項）を確認すること。

なお、確認に当たって所定の書類を提出しない場合も、賃上げ基準に達していない者と同様の措置を行う。

また、受託者は、経年的に賃上げ表明を行う場合、事業年度単位か暦年単位かの選択を前年度又は前年から変えることによって、前年度等に加点を受けるために表明した期間と、当該年度等に加点を受けるために表明した期間が重なり、賃上げ表明の期間と加点を受ける期間との間に不整合が生じることとなるため、賃上げ表明を行う期間は、前年度に加点を受けるために表明した期間と重ならないようにしなければならないことに留意すること。

## 12. その他

本委託業務に伴う知的財産は当研究所に帰属するものとする。ただし、成果物を作成するために受託者が開発したツール、本業務のために受託者が提供した知的財産の権利及び受託者が既に有していた著作権については受託者に権利が留保される。

この仕様書に記載されていない事項、または本仕様書について疑義が生じた場合は、科政研と適宜協議を行うものとする。また、本委託業務の実施にあたっては、会計に関する法令に定めるほか、科政研委託契約事務処理要領により適切に実施しなければならない。