

# 仕様書

## 1. 件名

J Dream IIIデータの抽出および整理 2024

## 2. 目的

科学技術・学術政策研究所(以下、「当研究所」という。)では、地域に自律的・持続的なイノベーションシステムが構築されるよう、地域における科学技術の基盤と活動の状況を解析することにより、地域科学技術に関する調査研究を実施している。

当研究所が行う調査研究において、大学における研究成果等の発表状況に係るデータ分析を実施する必要があることから、J Dream IIIのJSTPlusファイルに収録されている全86国立大学法人所属の研究者による研究成果等の発表データを抽出・整理することを目的とする。

## 3. 仕様

J Dream IIIのJSTPlusファイルに収録されている全86国立大学法人所属の研究者による研究成果等の発表データを、別添 注意事項に留意し、以下のとおり全86国立大学ごとに抽出・整理すること。

### ①対象期間:

2017年度～2022年度

### ②対象情報:

J Dream IIIのJSTPlusファイルに収録されている情報であって、上記期間内に全86国立大学法人所属の研究者が発表を行った文献に係る情報。(約130万件)

### ③データフォーマット:

上記対象情報を、全86国立大学ごとに抽出・整理しデータシートにすること。データシートのフォーマットはExcel形式とし、Microsoft Office Excel 2019以降で正しく表示されること。なお、Excel形式での提供にそぐわないものについては担当者との調整のうえ、作成すること。

### ④データの提供方法:

上記①～③に従い、データを抽出・整理し、当研究所の必要に応じて提供すること。なお、J Dream III利用約款における学術・病院向け固定料金サービス利用約款に記載されている第12条を遵守し、当研究所に職員を派遣させるなどして、保存場所と同時保存上限に留意し作業をすすめること。

※利用約款：[https://jdream3.com/guide/catalog/document/jd\\_price\\_agreement\\_2018.pdf](https://jdream3.com/guide/catalog/document/jd_price_agreement_2018.pdf)

### ⑤データ精度の担保:

データ抽出、整理作業において、ヒューマンエラーを回避することを目的とし、ダブルチェックを行うなど、データの精度を保つ対策を講じること。

4. 納入品
  3. 仕様に記載しているデータセットの全て
  
5. 無償貸与を行う資料  
当研究所が許諾を受けている J Dream III の ID、およびパスワード
  
6. 納入場所および履行期限等
  - (1) 納入場所  
文部科学省 科学技術・学術政策研究所 第2調査研究グループ  
〒100-0013 東京都千代田区霞が関3丁目2番2号  
中央合同庁舎第7号館東館16階
  - (2) 履行期限  
令和6年9月13日（金）
  
7. 検査及び検収
  - (1) 検査は、納入品等について、担当者が所定の検査を行い、これに合格したことをもって検収とする。
  - (2) 検査の結果、要求要件が満たされていない場合には、当研究所は請負者に対して修正等を求める。
  - (3) 請負者は、当研究所から修正等の指摘を受けた場合には、責任を持って速やかにこれに対応すること。
  
8. 保証期間  
請負者は、当研究所の責めによらない理由によって、当該データに欠陥等があった場合は、納入から1年間の期間内において無償で修復・交換をすること。
  
9. その他  
本仕様に定める事項に関して疑義を生じた場合は、当研究所担当職員と協議して指示を受けるものとする。

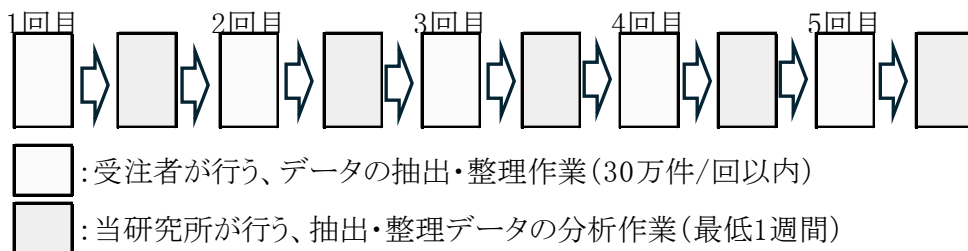
以上

(別添)

## 注意事項

### 1. データ抽出について

- ✓ **J Dream III**のJSTPlus ファイルから抽出される、データ1件当たりの項目  
「整理番号、和文標題、英文標題、原文標題、著者名、資料名、JST 資料番号、ISSN、CODEN、ISBN、レポート番号、巻号ページ、特殊号、発行年月日、写図表参、会議、資料種別、記事区分、発行国、言語、抄録、英文抄録、分類コード、シソーラス用語、準シソーラス用語、IPC(機械付与)、物質索引、著者名のみ、所属機関名のみ、著者名(所属機関)個別、所属機関名(機関ID)、機関ID、著者ID、DOI 情報、MyCollection1、MyCollection2、MyCollection3、引用数、引用テキスト(引用整理番号)、引用整理番号、被引用数、被引用テキスト(被引用整理番号)、被引用整理番号、著者名(著者ID)所属機関名(機関ID)、**JDreamIII 複写可否、ダウンロード日、COPYRIGHT**」
- ✓ ファイルに収録されている情報であって、全 86 国立大学法人所属の研究者が学会発表を行った文献を漏れがないように担当者と協議のうえ収集すること
- ✓ 著者により所属機関の表記が多岐にわたるため、表記揺れを理解し抽出すること
- ✓ 一度にダウンロードできるデータ量は上限が1万件のため、分割してダウンロードした後に結合すること
- ✓ 上記結合の際に漏れがないように注意して作業すること
- ✓ データの抽出精度により分析結果が異なるため、データセットについては責任を負うこと
- ✓ 利用約款の「第12条 3. 機械可読の形態で同時に保存できる文献数は各データベースごとに300,000 件を超えてはならない。」と記載されているように、保存できるデータ量に限りがあるため、当所で分析を行う期間を考慮し、担当者と協議のうえ納品スケジュールをたてること
- ✓ 1回の分析には最低 1 週間程度かかることを把握のうえ業務を行うこと  
想定作業手順(案)



\*データ抽出・整理作業が5回で終了するとは限らない。

### 2. データ整理について

- ✓ 分析が不能になるため、一つの大学法人のデータセットはデータ保存上限 300,000 件の中に必ず収まるように調整し、2 つのデータセットに跨らない様にすること

### 3. その他

- ✓ 担当者はマイクロソフト オフィス スペシャリスト(MOS)等の資格を有することが望ましい
- ✓ 学術論文データベースからデータの検索・抽出を行った経験がある方が担当するのが望ましい
- ✓ 書誌情報の整理を行った経験がある方が担当するのが望ましい
- ✓ 担当者は原則として操作マニュアルを自身で読解し、対応できることが望ましい