

特許情報取得APIの活用

Utilization of Patent Information Retrieval API

特許庁 総務部総務課特許情報室長

蛭田 敦

平成14年特許庁入庁。化学分野の審査・審判に従事。審判課企画班、特許審査第三部審査調査室、調整課審査基準室、総務課法規班、審判企画室、調整課企画調査班等を経て、令和3年10月より現職。

1 はじめに

特許庁は、令和4年1月から特許情報取得APIの試行提供を実施している。APIとは、「Application Programming Interface」の略語であり、ソフトウェア間を連携させる仕組みである。特許情報取得APIの利用者は、特許庁が保有する情報提供用のサーバに、利用者が保有するPC等をオンラインで接続することにより、様々な特許情報を取得することができる。

特許情報取得APIの提供は試行の段階であり、今後、利用者からの声に耳を傾けつつ、同APIの利用条件等を調整していく必要がある。本稿では、特許情報取得APIの提供に至るまでの過程や、同APIの特徴等を紹介したい。

2 API 提供に関する背景

政府は、デジタル社会の実現に向けた重点計画¹を定め、「デジタルの活用により、一人ひとりのニーズに合ったサービスを選ぶことができ、多様な幸せが実現できる社会の実現」に向けた各種施策に取り組んでいる。その中で、APIは、行政機関同士や行政機関と民間事業者等との情報交換を効率化・リアルタイム化させる手段として期待されており、API連携も含めたプラットフォーム整備などにより、行政機関と民間事業者等との連携や協力を積極的に推進する旨も盛り込まれている。また、経済産業省デジタル・ガバメント中長期計画においても、特許庁における業務・システム最適化の取組の1つとし

て、API提供の検討が掲げられている。

また、海外に目を向けると、複数の知的財産庁がAPIの提供を行っている。主要国・地域におけるAPIの提供状況は、令和元年度「特許情報提供サービスの現状と今後に関する調査」の報告書²で紹介されているが、その後APIの増加等のサービス拡大も見られる。

特許情報をAPIで提供するに当たっては、利用者のニーズに即した情報提供を行う必要があり、また、全ての特許情報を対象としてAPIを提供すると開発期間や運用コストが増大することから、特許庁では特許に関する情報のみを対象としてAPI開発を行うこととした。APIの開発は、アジャイル型と呼ばれる開発手法を採用することにより、利用者の意見を反映しやすい形とした。そして、APIで提供する情報の範囲等は、機械的な情報取得が可能になるというAPIの特性と、特許情報の活用形態とを考慮したものとした。

3 特許情報取得APIのコンセプト

特許庁が保有する特許情報は、工業所有権情報・研修館（INPIT）が運営するJ-PlatPatで提供されている。特許庁が発行した公報に関する情報のみならず、出願・審査に関する経過を示す情報、特許権等が登録された後の情報、そして、出願人から特許庁に提出された書類や、特許庁から出願人に発送された書類も確認することができる。これらの情報は、様々な利用者が特許情報検索・照会の経験によらず利用できるように、直感的に使いやすいインターフェイスを介して提供されている。

J-PlatPat の難点は、様々な利用者層に対して同一のインターフェイスでしか情報を提供できないことである。年に 1 億回以上のアクセスがある J-PlatPat の利用者は様々な属性を有しており、それぞれの利用者が利用しやすいと考える画面デザインや機能も一様ではない。ウェブサイトやアプリの宣伝で「さらに使いやすくなった」という宣伝を目にすることがあるが、一般論として使いやすくなったと評価される場合であっても、経験や慣れによって使いにくくなったと感じるユーザーがいるのが実情ではないだろうか。J-PlatPat に対する意見や要望は多く寄せられており、特許庁と INPIT とは改善に努めているものの、全ての利用者が納得するサービスとするのは容易ではない。

特許情報取得 API は、J-PlatPat で提供されている情報の範囲を超えて情報提供するものではないが、新たな情報の利用形態を提供するものである。例えば、J-PlatPat の利用者が特定の出願の経過情報を取得する場合、利用者は、インターフェイス上で出願番号を入力し、表示された情報をダウンロードするか書き出すことにより情報を取得する。API は、このような人手による作業を機械的に実現することを可能とするものであり、IT スキルを有する者であれば好みのタイミングで情報を取得し、取得した情報を自作したインターフェイス上に表示することができる。つまり、特許情報取得 API は、利用者の協力の下に、情報取得の効率化や多様なニーズに応えることを目的として提供されるものである。

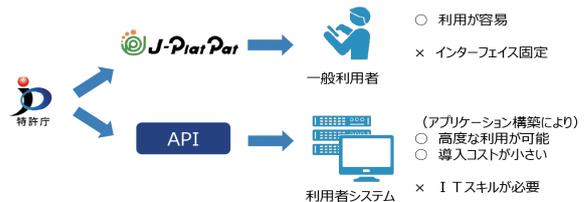


図 1 J-PlatPat と特許情報取得 API との比較

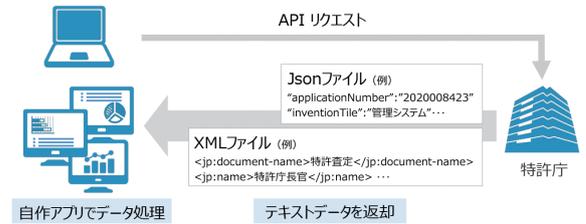


図 2 特許情報取得 API による情報取得

4 特許情報取得 API のコンテンツ

特許情報取得 API が提供する情報は、J-PlatPat で提供されている情報のうち、特許に関するものであって、主に、出願の経過情報や書類情報である。経過情報は、J-PlatPat で提供されているものの、情報取得に手間がかかることから、機械的に取得することが要望されていた。また、書類情報は、公報情報とは異なり民間事業者に対してバルク提供されていないために、民間事業者のデータベースからは提供できていなかった。

API を試行的に提供するに当たり、まずは表 1 に示す API を作成して提供することとした。特にニーズが

表 1 特許情報取得 API の種類

	名称	概要	アクセス上限	データ分類
1	特許経過情報	出願番号に基づき経過情報の一覧を取得	400	経過情報
2	シンプル版特許経過情報	1 の簡易版（優先権情報、分割情報を含まない）	400	経過情報
3	特許分割出願情報	出願番号に基づき分割出願情報を取得	30	その他
4	特許優先基礎出願情報	出願番号に基づき優先基礎出願情報を取得	30	その他
5	特許申請人氏名・名称	申請人コードに基づき申請人氏名・名称を取得	200	その他
6	特許申請人コード	申請人氏名・名称に基づき申請人コードを取得	200	その他
7	特許番号参照	出願 / 公開 / 登録番号に基づき相互に番号を取得	50	その他
8	特許申請書類	出願番号に基づき特許申請書類の実体ファイルを取得	100	書類情報
9	特許発送書類	出願番号に基づき特許発送書類の実体ファイルを取得	100	書類情報
10	特許拒絶理由通知書	出願番号に基づき拒絶理由通知書を取得	100	書類情報
11	特許引用文献情報	出願番号に基づき拒絶理由の引用文献情報を取得	50	その他
12	特許登録情報	出願番号に基づき登録情報を取得	200	経過情報

高い経過情報については、全ての経過情報を網羅的に提供する完全版と、完全版から優先権情報や分割情報を除外したシンプル版とを用意した。また、書類情報については、特許庁から発送される全ての書類を集めた特許発送書類 API だけでなく、特にニーズが高いと考えられる拒絶理由通知書に提供範囲を絞った特許拒絶理由通知 API も作成して提供することとした。その他に、出願人や代理人が特許庁へ手続を行う際に有用と考えられる申請人コード情報等、サポート情報に関する API も作成された。そして、情報へのアクセス頻度を考慮し、API へのアクセスについては、API の種類ごとに 1 日当たりの制限を設けることとした。

5 試行提供の状況

特許情報取得 API は、令和 4 年 1 月から試行提供が行われている。提供開始に当たり、事前説明会の実施とともに、関連情報を集めた情報提供サイト³の運用が開始され、利用条件や API の仕様に関する周知が行われた。特許情報取得 API の利用に際して必要な情報は、今後も情報提供サイトで継続提供される予定である。

API の利用希望者は、利用者登録申請を行い、申請内容が特許庁担当者により確認された後に API の利用を開始することができる。利用登録が完了した者に対して、特許庁担当者は、API 利用時の認証手続に必要な ID とパスワードを発行する。認証手続に成功した利用者に対しては、アクセストークンと呼ばれる認証キーが発行され、有効期限内でのアクセスが可能となる。

試行期間においては、一時的なアクセス過多による応答エラーの発生等はあるものの、API 提供を停止しなければならないような大きなトラブルは発生していない。(1) 利用者、(2) 利用時間帯、(3) 利用態様の観点で取りまとめた試行約半年間での利用状況は、次のとおりである。

(1) 利用者

利用登録者は、試行開始から半年で約 200 者となった。利用登録者の属性別では、一般企業による登録が最も多く、次に個人による登録が多い。また、DB 事業者(特許情報の提供をサービスとしている民間事業者)や、特許事務所(弁理士事務所)による登録がそれぞれ 1 割

程度となっている。

他方、API のアクセスを見ると、一般企業からの利用割合が最も多い点は登録状況と同様であるが、個人によるアクセスは登録割合に比して少なく、特許事務所からのアクセスが多い。特許事務所では、事務所が保有するデータベースへのデータ入力に特許情報取得 API を活用するなど、日常的にアクセスする仕組みを設けているところが複数あり、その結果として、アクセス割合が増えていると考えられる。個人登録者による利用は、API 利用に必要なアプリケーションの準備等に相応の時間を要することなどから、進んでいないものと考えられる。

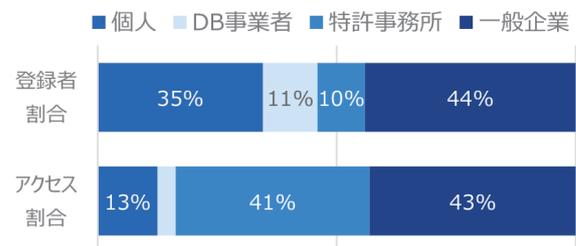


図3 特許情報取得 API の利用者の傾向

API の試行提供環境の制限から、法人は、複数の ID 登録を行うことができず、従業員が ID を共有する形で利用することとされている。DB 事業者からは、商用サービスに API を適用して機能拡充を検討しているものの、アクセス可能数が少なすぎるために、API を利用することができないとの声もある。後述するように、DB 事業者と連携して特許情報取得 API を提供することにより、通信量を制御しつつ効率的な情報提供ができる可能性もあるため、利用者の属性に応じたアクセス上限を設定するなど、より柔軟な利用条件の設定を行うことで、利用者が利用しやすいサービスとすることが課題である。

(2) 利用時間帯

特許情報取得 API の特徴は、原則として、24 時間 365 日利用することができ、利用者の業務時間外であっても機械的に情報を取得することが可能な点にある。API を提供する特許庁としては、特定の時間帯にアクセスが集中し、利用者からのリクエストに適時に応えられない事象が発生しないかが懸念点である。その一方で、アクセス集中を見越したハードウェア環境を整備すると、過剰な設備投資になるおそれもある。そこで、

利用者には、アクセスが集中しにくい夜間や土日祝日での利用を推奨している。上述のとおり、APIにはアクセス制限があるところ、夜間や土日祝日での利用に協力可能な法人に対しては、アクセス制限を緩和することで、アクセスの分散が図られている（図4参照）。

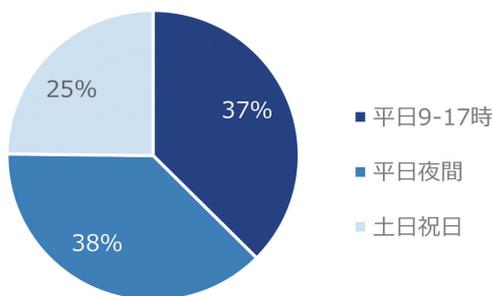


図4 特許情報取得APIの利用時間

利用者には、引き続き、APIの用途を踏まえて、アクセスを分散化させる利用に協力していただきたい。特許情報取得APIで提供される情報は平日であれば毎日更新されるため、例えば、始業時間までに必要な情報をAPIで自動取得し、利用者が保有するデータベースにデータ入力を機械的に行っておくことで、業務時間における利用者のデータベースの情報を最新に保つというような利用方法が考えられる。

(3) 利用形態

特許情報取得APIで得られる情報の利用方法は、利用者によって様々である。APIの種類（データ分類）ごとのアクセスを見ると、経過情報へのアクセスが最も多く、次に書類情報が参照されている。

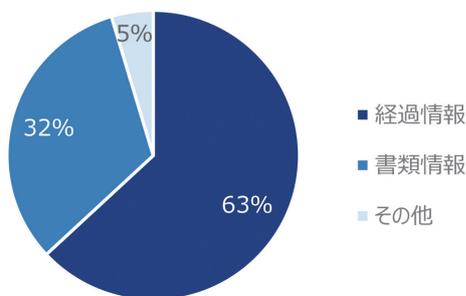


図5 データ分類別のアクセス量

経過情報に関するAPIは、例えば一般企業であれば、自社出願の管理に利用されているケースが多く、社内

データベースへ情報を取り込んで利用されていることが多い。特許事務所は、手続代理を依頼された出願の管理のみならず、注目している特許出願の状態の監視にも利用している。ここで、J-PlatPatで提供されているRSS機能では出願状態の更新を通知しているため、RSS機能とAPIとを組み合わせることで、情報更新があったときだけAPIで情報取得するというように、更に効率的な情報取得が可能になると考えられる。RSS機能の利用方法は、J-PlatPat操作説明動画⁴により確認することができる。

書類情報に関するAPIは、書類データを利用者が保有するデータベースに格納するために利用されているケースが多く、その上で、拒絶理由分析等の分析にも役立てられている。書類情報は、複数の書類がまとめられてZIP形式で提供されるが、それぞれの書類はXML形式の構造化されたファイルで提供される。XML形式のファイルは機械的な分析を行いやすく、J-PlatPatで提供されている書類情報からは取得できない情報でもあるため、特に拒絶理由分析等に興味のある利用者は、書類情報APIを積極的に活用して欲しい。

特許情報取得APIの試行期間は1年間とし、利用が少ない場合には廃止することも視野に入っていたが、一定のニーズが確認されたことから、API提供は、今後も継続される予定となっている。また、現在は試行提供中であるため、アクセス上限が厳しめに設定されているが、利用状況を注視しつつ、徐々にアクセス制限を緩和させていく予定である。

特許情報取得APIの試行は、利用者からのサポートにも支えられてきた。例えば、試行開始初期においては、利用時の認証手続に関する課題が発覚したが、利用者からのインターネットでの情報発信等によって、課題が解消され、特許庁への問合せも大幅に減少した。上述のとおり、アクセス量の平準化なども含め、特許情報取得APIの運用において、利用者の協力を支えられていることに感謝を申し上げたい。

6 特許情報取得APIの今後

特許情報取得APIの活用状況からみて、利用者は今後さらに増加し、活用の幅も広がっていくと考えられる。さらなる活用を促すためには、アクセス制限を緩和する

こと、そして、APIの種類を増加を行うことが必要である。私案ではあるが、今後の特許情報提供 API のあり方について述べたい。

現在の試行段階においては、アクセス制限があるために、制限内で業務に活用できる利用者に利用が限定されているおそれがある。特許情報の提供において、特許庁は正確で基本的な一次情報を提供し、民間が付加価値の付いた二次情報を提供するという方針が産業構造審議会等で確認されている⁵。そして、同審議会の報告では、「民間事業者が高い付加価値を付けたサービスを提供していくための環境を整備し、我が国ユーザーによる一層高度な特許情報の利用を促していくべきである」とも提言されている。上述のとおり、データベースを提供する民間事業者が特許情報取得 API を有効に活用できているとはいえない状況であるため、民間事業者にとっても利用しやすい API 環境を目指すべきであると考えられる。

民間事業者にとって、特に書類情報に関する API は、利用価値の高いものであると考えられる。民間事業者のサービスの多くは、特許庁が発行する公報に関するデータや、出願の状態に関する諸データを集めた特許情報標準データを利用して展開されている。仮に、書類情報を民間事業者のサービスから提供する場合、これまでの情報共有の形を踏襲するならば、特許庁は、日次で書類情報をバルク提供し、民間事業者は、バルク提供された書類情報を日次で自社データベースに反映させなければならない。そして、データの蓄積漏れはデータベースとしての欠陥となり得るため、民間事業者は、全ての書類情報を蓄積するために必要なサーバ等を準備する必要がある。それに対して、API を利用する場合、書類情報を蓄積するためのサーバを用意する必要はなく、民間事業者は、クライアントからのリクエストに応じて API を通じた情報取得を行う仕組みを構築するだけで良い(図6参照)。API を利用することは、民間事業者のサーバ等への設備投資を抑え、国全体として、特許情報提供に要するコストを下げることにつながると考えられる。

特許庁が提供する API に民間事業者のサービスが依存するようになると、さらに当該 API を安定稼働させる必要性が高まることになる。特許庁としては、信頼される API サービスを提供できるように、試行期間中に得られた知見を活用していく所存である。

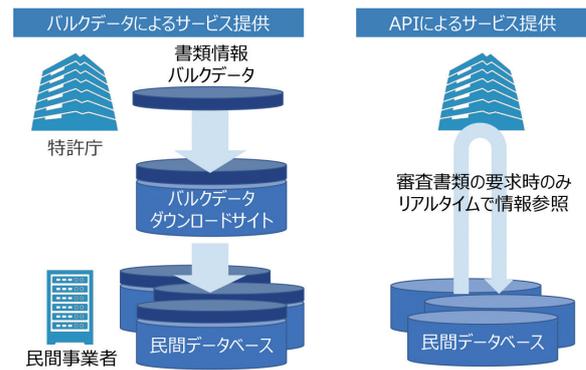


図6 民間事業者による書類情報提供のイメージ

ここでは、民間事業者による特許情報取得 API の活用案について述べたが、その他の利用者による活用の拡大も期待したい。利用者間でのコミュニケーションの輪が広がることで、さらなる活用方法が発見されることも期待される。API の利用者が増加し、利用者間での情報共有が進み、新たな活用形態が生まれ、さらに多くの利用者を呼び込むという好循環により、特許情報の普及が進めば幸いである。

7 その他の特許情報施策

筆者が所属する特許情報室は、特許情報の普及、活用の促進のために、ユーザーニーズの調査を行いつつ、様々な施策を行っている。令和3年度末には、特許情報提供サービスの現状と今後に関する調査報告書に加え、機械翻訳に関する複数の調査報告書⁶が取りまとめられ、特許庁のウェブサイトに掲載された。そこで、最後に、これらの調査報告書について簡単に紹介させていただく。

(1) 特許情報提供サービスの現状と今後に関する調査

特許情報提供サービスの現状と今後について、令和2年度の調査は社会情勢を踏まえて見送られたが、令和3年度の調査は実施されている。当該調査においては、海外の主要な知的財産庁が提供するサービスや、特許情報利用者による海外の特許情報の取得方法等について調査が行われた。

(2) 機械翻訳に関する調査報告書

日本の特許情報の海外への発信や、海外の特許情報

の日本ユーザーへの提供において機械翻訳が活用されているところ、機械翻訳に関する調査事業が定期的に行われている。令和3年度には、機械翻訳エンジンの学習に用いられるデータ、すなわち、原文と翻訳文との文対（コーパス）データが、翻訳品質に及ぼす効果を調査した3つの事業で報告書が取りまとめられた。そのうち、審査書類・審決の機械翻訳（日本語→英語）と、中国審決の機械翻訳（中国語→日本語）に関する2つの調査は、それぞれ3年をかけて実施したものであり、コーパスの学習により翻訳品質が向上することが確認された。中国公報の機械翻訳（中国語→日本語）において機械翻訳できなかった用語に関する調査では、新たな技術用語等を機械翻訳エンジンに辞書登録することにより翻訳品質の向上が確認された一方で、機械翻訳前の文章解析（文分割等）が適切に行われない場合に誤訳や翻訳漏れが生じる傾向にあることが分かった。

機械翻訳に関する調査で得られた結果や、調査の過程で作成したコーパスは、特許庁の機械翻訳プラットフォームの運用や、外国特許情報サービス FOPISER⁷ における中国審決の和訳文照会サービス等で活用される予定である。

特許情報に関する調査事業は、令和4年度も実施し、特許情報の利用者の意見等を今後の施策に反映させる予定である。特許情報の利用者には、当該調査事業への協力もお願いしたい。

参考文献

1. 「デジタル社会の実現に向けた重点計画」、令和4年6月
https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/5ecac8cc-50f1-4168-b989-2bcaabffe870/d130556b/20220607_policies_priority_outline_05.pdf
2. 特許情報提供サービスに関する調査報告書について
<https://www.jpo.go.jp/resources/report/sonota/service/index.html>
3. API 情報提供サイト
<https://ip-data.jpo.go.jp/pages/top.html>
4. ◇経過情報のウォッチングについて（RSS 機能）
https://www.inpit.go.jp/jinzai/ipeplat/info_20220124.html
5. 産業構造審議会 知的財産分科会 情報普及・活用小委員会報告書「特許情報のさらなる活用に向けて」、平成28年5月
6. 機械翻訳に関する調査報告書
https://www.jpo.go.jp/system/laws/sesaku/kikaihonyaku/kikai_honyaku.html
7. 外国特許情報サービス FOPISER
<https://www.foreignsearch2.jpo.go.jp/>