

慶應義塾大学におけるグーグル・ライブラリー・プロジェクトの著作権調査について

さとう ゆりえ
佐藤友里恵

(三田メディアセンター係主任)

1 はじめに

2007年7月に慶應義塾創立150年記念事業の一環としてグーグル社と連携したプロジェクトがスタートした。このプロジェクトは、グーグル社と大学図書館が連携して蔵書をデジタル化するライブラリー・プロジェクトである。慶應義塾大学は、欧米地域以外で初めてこのプロジェクトに参加して、三田メディアセンターで所蔵する資料のうち著作権保護期間が満了しているものをデジタル化することになった。

まず第一フェーズとして著作権の問題のない慶應義塾と福澤諭吉に関係する書籍173冊のデジタル化を行い、慶應義塾創立150年の年である2008年1月10日の福澤先生誕生記念会に合わせて公開した。三田メディアセンターでは併せてGoogleブック検索のデモンストレーションも行った。

第二フェーズは一般蔵書のデジタル化で、事前に著作権調査をはじめ、書誌・所蔵のメタデータ作成等の準備に多くの時間が必要で、実際にスキニング作業が開始したのは2009年5月であった。

この度の報告では準備作業の中でも筆者が2008年6月より関わってきた著作権調査について述べたいと思う。

2 著作権調査とは

今回のプロジェクトでは慶應に著作権のある資料はできるだけデジタル化して公開することを目標に、所蔵調査を行い、現物1点1点を確認した。次に一般蔵書をデジタル化して公開するためには著作権保護の対象とならないこと、あるいは著作権保護期間が満了していることを調査しなくてはならない。

ここで関係する著作権の種類は、大きく分けると「個人著作者の著作権」と「団体著作者の著作権」の2種類である。個人の著作権は没後50年経過で満了、団体の著作権は公表後50年経過で満了となる。このプロジェクトは当初2009年に公開することを前提としていたため、個人は没年が1958年以前、団体は刊行年が1958年以前のを公開の条件とし

た。また、1941年以降に刊行された個人著作物は著作権をクリアする確率が低いと想定し、作業効率を考慮して予め対象外とした。

その結果、第二フェーズではある程度まとまった単位で出せる資料群として、旧分類和図書(1961年まで使用していた請求記号体系のもの)、個人文庫の一部の和図書、和統計資料、和雑誌資料、1889年までに刊行された和装本が対象に選定された。

3 著作権調査チームの立ち上げ

調査を集中して行うために、2007年9月より1年計画でプロジェクト体制をとり派遣職員3名+専任職員1名で作業を進めることになった(派遣は開始当初は1名、翌10月より3名の体制となった)。しかし予想以上に調査が進まず、予定期限までに終了できる見込みがたなくなり、同年12月よりレファレンス担当、2008年1月より雑誌担当がバックアップに入る形をとった。さらに調査のために必要な人名記入のある頁のコピーをとる作業については閲覧担当の協力も得られた。

このように周囲の援助を受けつつ調査を行ったが1年では終わらず、派遣の契約を1名は2008年12月末まで、残る2名は2009年1月末まで更新して期間延長で作業を進めた。それ以降は専任1名で残る処理を行ってきたものの、プロジェクト全体の日程が決定できないこともあり、調査途中のものを残して2009年10月末で著作権調査は終了した。

また、著作権調査を行うにあたり、国立国会図書館の近代デジタルライブラリー(以下、近デジ)構築のための「国立国会図書館資料デジタル化の手引き」を参考とした。調査対象の洗い出しや調査方法は取り入れたが、保護期間中あるいは保護期間中かどうか不明だった場合は、許諾処理や文化庁長官の裁定申請を行うことまではしなかった。時間と手間との制約があり、慶應独自の作業では困難と判断してやむを得ず省略したためである。

4 実際の著作権調査作業

著作権調査を行ったのは主に旧分類和図書で、その概要は以下のとおりである。次頁の流れ図を参照していただきたい。

処理の中で得られた情報（生没年、証拠、その他の情報、公開の可否）は、著作権調査を行うために独自に開発したデータベース（以下、著作権データベース）に入力し、管理している。このデータベースでは書誌レコードと著者レコードを作成して相互にリンクをはっている。また本名やペンネーム、号などでもそれぞれレコードを作成してリンクをはり、調査の手間を省いた。

(1) 著者拾い出し

最初に発行年を確認し、1941年以降であれば調査不要として対象から外した。

次に資料の状態を確認し、折込図のみからなるもの、厚みがあって頁が開きにくいもの、10頁以上の水濡れがあるもの、大きく破損しているものは、これも調査不要で非公開とした。

それから近デジを検索し、生没年が明確なものはそのまま利用して、それ以外（許諾・裁定のもの）は参考情報とした。

次に資料の内容を見て、公開できない個人情報がかかっている等公開には相応しくないと判断されたもの、拾い出す対象者が11名以上で調査に時間がかかるもの、中国や朝鮮で発行された漢籍、すでに著作権保護期間中の著者がいると確認できたものは、この段階で非公開とした。

以上の事前処理を行ってから著者の拾い出しにかかった。見るポイントは、標題紙・奥付・背・表紙の著者名、序文・前書き・目次・後書きの内容、挿絵・歌の作者名やサイン、本文中の文章のかたまり（章や論文単位）ごとの著者名である。拾い出し中に保護期間中と判明した著者が出てきた場合には、非公開が確定なので、その時点で調査を中止した。

(2) 書誌レコードの処理

リンクしている著者の没年がすべて判明していて、かつ保護期間が満了であった場合は書誌ステータスを「公開」として、1人でも保護期間中の著者がいれば「非公開」として調査を終了した。

没年不明の著者がいるが、判明している内の1人でも保護期間中であれば「非公開」として調査を終了し、不明の著者は次回調査へ回し、次に出てきた

時に改めて調査を行うこととした。

没年不明の著者がいて、かつ保護期間中の人がいなければ「調査中」として没年調査の結果を待ち、最終的に「公開」または「非公開」に振り分けた。

(3) 著者レコードの処理

拾い出した著者について、まずは著者レコードの有無と書誌レコードへのリンクを確認した。リンクしている著者以外にも拾い出し対象者がいれば、既存のレコードを検索して書誌レコードとリンクするか、新規に作成してリンクした。

次に著者が個人か団体かを確認し、団体であれば保護期間満了として終了である。個人の没年が入っていれば、ステータスに「判明」、著作権クリアの状態を「満了」あるいは「保護期間中」とした。

没年が入っていなければ「調査中」として没年調査へ回し、以下の(4)～(6)の結果を待って最終的に「判明」か、これ以上の調査は困難として「調査不可能」に振り分けた。

(4) 第一次没年調査

ここでの没年調査のツールには、WHOPLUS、近デジ、CiNiiなど、机上で利用できるデータベースを用いた。

没年が判明すれば、その年とツール名、ステータスを入力する。判明しなくても近デジで満了だった場合は満了の扱いとした。

没年が判明していなくても1958年以前に死亡していること、あるいは1959年の時点で生存していたことが明確な場合には、それぞれ判明したものとして扱った。この場合は根拠となった情報をツール欄に入力し、コピーもファイルしておいた。

判明しなかったものは「調査中」として、参考情報とともに次の調査へと回された。

(5) 第二次没年調査

ここでは、皓星社の『データベース版・雑誌記事索引集成』（日本人物情報大系）、Googleブック検索、県別人名辞典類、慶應関係レファレンス資料、ほか主題に合わせたデータベースやレファレンス資料を用いて調査をした。

没年が判明したら著作権調査チームに渡され、著作権データベースに入力された。判明しなかったものは、それまでに得られた情報すべてを添えて最終調査へと回された。

(6) 第三次没年調査

添付された情報と資料の内容全体からの情報を元に、三田メディアセンター内の蔵書、あるいは他キャンパス、他大学図書館、関連機関へのILLを通して調査を進めた。現物取り寄せは学内と早稲田大学図書館のみにとどめた。

没年が判明したら著作権調査チームに渡され、著作権データベースに入力された。判明しなかったものは「調査不可能」として最終的に調査中止とした。

「保護期間中」あるいは「調査不可能」と判断された時点で、この著者がリンクしている書誌をすべて先に「調査不要」(=非公開)とすることで、さらに効率的な作業を図った。

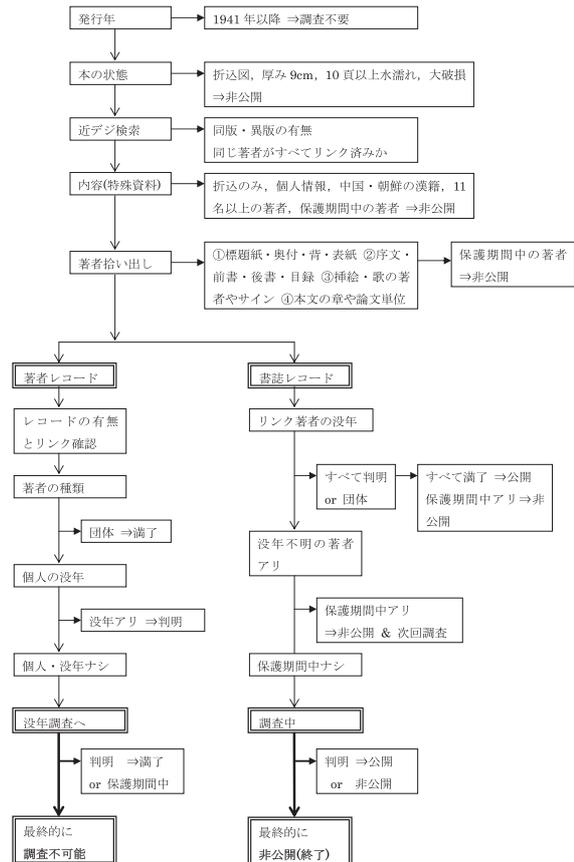
(7) 調査結果

作業に関わる者は著作権の専門家ではないので、解釈に迷うケースについては、専門の教員などからアドバイスを受けつつ処理方針を決めていった。

また、処理の中で出された質問をその都度Q&Aのファイルに入力し、類似のケースや前後する時間の中で解釈・判断がぶれないように注意を重ねた。

このようにして一通りの調査を終えた結果、団体を含めた全著者の61.5%が判明し、満了と判断された著者は40.5%であった。また、判明した数に対する満了の率は65.9%であった。

半数を超える著者が判明に至ったのは、慶應義塾の蔵書の幅広さと、レファレンス担当の力量によるものと思う。CiNiiやGoogleブック検索でヒントとなる情報が載っている文献が予想以上に多くヒットしたことにも驚いたが、かなりの率でそれらの資料を三田メディアセンターで所蔵していることにも感心させられた。



5 著作権調査での迷い

学術研究に貢献するため、できるだけ多くの蔵書を公開したいと考える。しかし、Googleブック検索が著作権者の許諾を得ずにデジタル化を行ったとして世界の大きな注目を浴び、その後和解論争が日本にも波及したことを考えると、慶應義塾が提供するコンテンツについて著者の権利を無視したというクレームはどうしても避けたい。著作権調査は重圧に耐えながらの作業であった。

著者拾い出しにおいても、拾うべき著者か否か迷ったり、拾ったとしても時代的にくずし字が多く判読できなかつたり、また挿絵のサインや落款が判読できなかつたりして第一歩から躓くものも少なかつた。判読困難なものについては貴重書室担当に大いに助けてもらった。また漢字や名前の読み方についても知識や経験が足りないと、一つ違うだけでヒントから遠ざかってしまうこともあった。

また、ほぼ同一人物と思われても確証がなければ判明には至らない。個人のHPなどでは没年が掲載されているのに、それが信頼できる情報源とはい言切れないために「調査不可能」となったもの、生年

から推測して満了となっている可能性が非常に高くても決め手に欠けるもの、1960年前後の他者の論文などから執筆された時点ではもう亡くなっていることがわかるが、それがいつなのか明瞭でないものなど、あと一歩というところで断念したケースも多い。もう少し追跡すれば判明するかもしれないと思いつつも時間的な制約もあり、どこで打ち切るか、特に最終段階の調査では悩みどころだったのではないかとと思われる。

6 プロジェクトの今後

2010年8月現在、スキヤニング作業も最終段階に入り和装本を処理中である。2011年1月にスキヤニングが終了する予定であり、今後は公開データのチェック、APIを使ったKOSMOS（慶應のOPAC名称）からのリンク等について検討に入ることになる。

Googleブック検索上にはスキヤニングされたイメージデータと、書誌のメタデータが整えられて公開へと至る。現在、旧分類和図書はかなりの数が公開されているが、メタデータ・画像・OCRの質等で問題のあるものもあり、グーグル社とともに検証中である。

このプロジェクト用に作成された書誌データは所蔵データを含んだものであり、旧分類和図書については1所蔵に対して1書誌が作成された完全な物理

単位レコードから構成されている。書誌を作成する際には重複チェックなどしていないので、当然のことながらデータベース内には書誌重複レコードが数多く存在する。また目録データベースから得られた書誌データではなくオリジナルで作成されたものは、目録担当者が処理したわけではないのでやはり質が劣る。遡及データとも言えるものなので蔵書検索に役立てたいのだが、一方でそのまま利用するのは逆にノイズを増やすことになるのではないかと懸念される。

また、イメージデータはグーグル社よりコピーをもらうことができる。グーグル社とは異なるインターフェースで公開することを予定しているが、どのように活用するのか、これについても検討していかなければならない。

蔵書を電子化して公開することについて、また、このプロジェクトの対象に選定された資料群についていろいろ意見もあるだろう。しかし、著作権の調査を進める上で、その「電子化された資料」にかなりの部分で助けられたのも事実であり、なかなか入手しにくくなった資料を閲覧できるという利点があるのは確かである。これだけの時間と手間をかけたものであるから、それを必要とする方々に大いに役立てていただきたいと願う。



本を護る：グーグルの取り組み

スキャンする図書を運搬するために、特別の容器が作られた。本を傷めないようクッションを入れ、水濡れしないように蓋にも工夫がなされている。今回プロジェクトに参加するにあたり、一番の課題は資料を傷めることがあってはならないということだった。資料の運搬・スキャン作業・本の取り扱いについては長い時間をかけて合意を形成した。この容器もグーグル社の特別製でトーツと呼んでいる。また利用者の不安の声に応えるため、グーグル社の了解を得て数名の教員の方に実際の作業の確認、スキャン前後での資料の状態をチェックしていただく態勢も整えた。書架に戻された本を見ていただければ、図書館員ばかりでなくグーグルのスタッフも資料の取り扱いには細心の注意を払っていることがわかっていただけたと思う。