

学術データベースとして、消費者サービスとして — 図書館用コンテンツ配信サービスの二面性

岩崎 陽一^{1,a)}

概要: 本稿は、ひとつの図書館用音楽コンテンツ配信サービスの事例を取りあげ、デジタル・アーカイブやデジタル・ライブラリーが人文学の研究の発展に寄与するために具えるべき情報検索インターフェイスと、それが依存するメタデータ構造について論じる。音楽コンテンツ配信サービスが学術データベースとして有効に機能するために設計した諸機能について、その設計思想を含めて抽象化して記述するという方法をとる。キーとなる発想は、資料の物理的形態に引きずられた設計を抜け出し、利用者の潜在的な要求を満たす方法を考え直すことにある。

キーワード: デジタル・アーカイブ, デジタル・ライブラリー, 情報検索, FRBR, Solr

To Be a Scholarly Database: Two Aspects of a Digital Content Distribution Services for Libraries

Abstract: This is a case study of a digital music distribution service for libraries to investigate into requirements for the information retrieval interface and metadata structure of digital archives and digital libraries to support the researches activities in humanities. The fundamental idea described here is that the interface should not be designed after physical forms of the content items.

Keywords: digital archive, digital library, information retrieval, FRBR, Solr

1. はじめに

人文学研究におけるデジタル・アーカイブやデジタル・ライブラリーの果たすべき役割や満たすべき要件について、既に多くの議論が行われている。しかし、それらにおける情報検索 (information retrieval) の機能要件やユーザー・インターフェイス設計についての議論は、特に国内では意外なほど少ない。その結果として、利用価値の高いコンテンツを効率的に提示できていないと、ユーザビリティの低いウェブサイトをししばしば目にするようになる。トップページに無数に並ぶ細かすぎる検索条件入力ボックスや、ソートも絞り込みもできない検索結果リストなどは、こんにち、学術サービス以外で新たに実装されることは稀である。資料の保存や管理の責任を負う MLA 職

員の視点から見れば、アーカイブやライブラリーの構築において、コンテンツのプレゼンテーションは優先度の高いものではないのかもしれない。しかし、デジタル・リソースの活用により人文学研究を推進しようという視点から見るととき *1、情報検索の利便性は決して軽視され得ない。総務省の 2003 年の白書においても、デジタル・アーカイブは文化の集積のみならず発信の拠点としても位置付けられている *2。貴重な資料がいかに精密にデジタル化され、大切に保存されていようとも、研究者がそのコンテンツに到達できなければ、研究推進という目標は達成できない。さらに、人文学研究者に対してデジタル・リソースの活用を促そうとするならば、「がんばって検索すれば何とか見つかる」という水準のユーザー・インターフェイスは望まれる結果をもたらさないだろう。そのためには、デジタル・リソースに不慣れな研究者が容易に使用でき、その研究を

¹ 東京大学大学院人文社会系研究科博士課程
Graduate School of Humanities and Sociology, University of Tokyo

^{a)} iwasakiyoichi@gmail.com

*1 Cf. 文献 [1].

*2 Cf. 文献 [2].

補助する資料を発見できるインターフェイスを実現しなければならない。

本稿は、学術資料を提供するデジタル・アーカイブおよびデジタル・ライブラリーにおいて、検索機能、およびそれが依存するメタデータ構造に求められる要件を検討する。論述は、Naxos Music Library*³ という音楽コンテンツ配信サービスのケース・スタディというかたちを取る。このサービスは、音楽愛好家に向けた消費者サービスと、図書館での利用を想定した学術データベースという、ふたつの側面をもっている。そのふたつのうち、学術データベースとしての役割を果たすために実装された諸機能を抽出することで、学術的サービスに求められる要件を提示できるだろう。本稿における要求分析の方法および実装方法は汎用性をもち、他のプロジェクトにも適用可能と考えられる。

なお、コンテンツのプレゼンテーションに関して、メタデータの可視化やアーカイブ間の相互運用については先進的な研究と実装が進められている。しかしこれらは必ずしも、アーカイブやライブラリーのウェブサイトにも求められる基本的な検索機能の利便性の向上に寄与するものでなく、付加的にはたらくものといえるであろう。それゆえ、本稿の考察の対象外とした。

2. Naxos Music Library の概要

Naxos Music Library (以下「NML」) は、Naxos Digital Services Ltd. が 2004 年に図書館市場に向けてリリースした、クラシック音楽を中心とする定額制デジタル音楽配信サービスである。翌 2005 年にはナクソス・デジタル・ジャパン株式会社 (現ナクソス・ジャパン株式会社、以下「ナクソス社」) が日本語版*⁴ を開発し、学術的電子資料として大学図書館、公共図書館で普及しているほか、音楽・放送業界のプロフェッショナル、および個人の音楽愛好家にも利用されている。本稿執筆時点 (2014 年 4 月 27 日) で約 700 レーベルの保有する、CD にして約 80,000 枚分の録音資料を所蔵しており、いまだ拡大を続けている。

筆者は NML 日本語版の立ち上げ時にナクソス社の社員としてその担当に就き、サービスと事業の開発に携わってきた。いちど同社を退職したが、2011 年より同社の技術顧問に着任。現在は、デジタル・ヒューマニティーズの動向を踏まえつつ、次世代の図書館の要求に叶う、或いはそれを先取りするサービスを提供するため、研究と提案を行っている。なお、本稿で紹介する試みの多くは日本語版でのみ実装している。本稿では、話題が日本語版に限定される場合は、サービスを「NML 日本語版」と明記する。

2004 年当時は珍しかった定額制音楽配信サービスも、いまや毎年いくつも新しいサービスが立ち上がり、そして消えていく。そのなかでほぼ唯一、NML 日本語版が日本国

内の図書館で学術データベースとして広く利用されている*⁵ 理由は、収録内容や販売方法だけでなく、それが図書館や研究者の要求に応えられるよう設計、開発されているということにもあるだろう。以下、そのような側面を形成する要素をいくつか紹介する。

3. データ提示方法

3.1 3 階層構造

まず、本稿の議論の前提として、NML の書誌データベースの特徴である 3 階層構造について説明する。商用音楽配信サービスのほとんどは、アルバムとトラックという 2 階層の書誌データ構造を採用している。これは、ひとつの作品の録音が原則として CD のひとつのトラックに記録される、ポップスやロックの書誌情報を想定すると妥当に思われるだろうが、NML が主に扱うクラシック音楽作品の録音物を記述するには適当でない。クラシック音楽作品の CD においては、通常、トラックはひとつの作品ではなく、それを構成する楽章やアリアに対応付けられる。例えば「ベートーヴェンの第九」という作品の演奏は、4 から 6 のトラックに分割して記録される。ひとつの作品に対応付けられる物質的な実体は存在しない。したがって、NML では作品の録音に対応する「トラックグループ」という概念を導入し、CD やアナログレコードの書誌データを「アルバム」—「トラックグループ」—「トラック」という 3 階層で記述している。クラシック音楽以外の作品においても、ひとつの作品を複数のトラックに分割して記録することは珍しくない。録音資料の書誌データを精密に扱うならば、この 3 階層構造は必須の要件となるだろう。なお、ここでは「アルバム」という概念を便宜的に用いる。これは「シングル」と対比されるものではなく、それが物質的商品として流通されるときにひとつの商品とみなされる実体を意味するものとする。

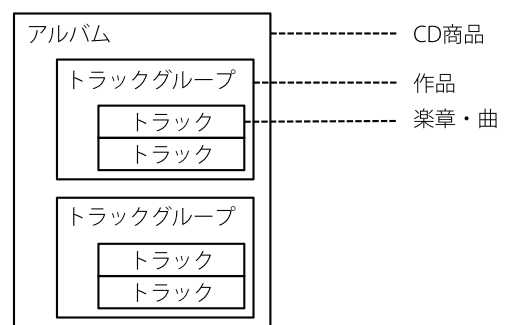


図 1 データ形式

複数のディスクで構成される「アルバム」においては、「ディスク」という概念も書誌データの単位となり得るが、

*⁵ 文献 [3] では音楽のオンライン資料としてナクソス社の他に 3 社のものが挙げられている。導入実績の統計資料は持ち合わせないが、いずれも国内で広く利用されているとはいいたい。

*³ <http://www.naxosmusiclibrary.com/>

*⁴ <http://ml.naxos.jp/>

NMLはそれを採用しない。オペラやバレエの作品では、複数のディスクにまたがってひとつの作品の録音が記録されることも多く、その場合、ディスクという単位は媒体の制約による分割点を示す以外に意味をなさないからである。

3.2 検索結果

OPACの類の書誌データベースでは、検索結果としてユーザーに示されるのは所蔵資料のリストである。録音資料の場合、多くは「アルバム」のリストがそれになる。一方、商用音楽配信では、ダウンロード販売の商品単位である「トラック」、ときにはそれに加えて「アルバム」のリストが検索結果として与えられることが多い。この仕様には、いずれも、商品流通や商品管理の事情が影響していると考えられる。図書館やデジタル配信業者にとって、仕入れて、管理し、流通させるべき実体は、商品単位である「アルバム」や「トラック」であり、自然とそれがカタログングの単位にもなる。しかしそれは自然ではあっても、必然ではない。

NML日本語版の設計では、常に、物質的形態に束縛された常識を離れ、デジタル・ライブラリーの理想的な姿を求めている。常識的に考えて、音楽配信のサイトであるならば、検索結果は「アルバム」か「トラック」のリストである。しかし、このサービスの利用目的を考えると、いずれも適切とは考えられない。クラシック音楽作品の検索において、利用者が特定の楽章のみを利用したいというケースは稀であろう。したがって、「トラック」のリストを返しても、重複行を無駄に増やすだけである。一方、「アルバム」を提示することも不適切と考えられる。たしかに、録音資料を「アルバム」単位で認識するユーザーは少なくない。どのレーベルから出ている、誰が何年にどこの交響楽団を指揮したCDが欲しい、というような条件で録音資料を探すことは、とくに音楽愛好家の間では一般的に行われている。それゆえ、音楽愛好家の要求を満たすよう、NMLでは「アルバム」の検索機能を提供している。しかし、NMLは利用者の誰もが音楽愛好家であるとは想定していない。むしろ、CDをあまり購入することのない人文系の研究者が、大学のインフラを通して手軽に音楽資料を利用し、それぞれの研究に取り入れてもらいたいと考えている。彼らがかしらの録音資料を参照しようとするとき——例えばインド学研究者がインド思想の西洋音楽での受容を調べるにあたってメシアン(Olivier Messiaen)の「トゥランガリラ交響曲(Turangalila Symphony)」*6を聴いてみようとするとき、どの「アルバム」の収録音源であるかということは重要ではない。利用者が求めるのは特定の「アルバム」でも「トラック」でもなく、特定の作品の任意の録音(「トラックグループ」)であろう。NMLでは、このような

検索に対応するため、「トラックグループ」を検索結果として表示する機能を有している。「トゥランガリラ交響曲」というワードでトラックグループを検索すれば、その録音がリストされる。利用者はそのなかで任意のもの、例えば演奏家やレーベルが有名なものや、リリース年代が新しいものを選んで利用する。

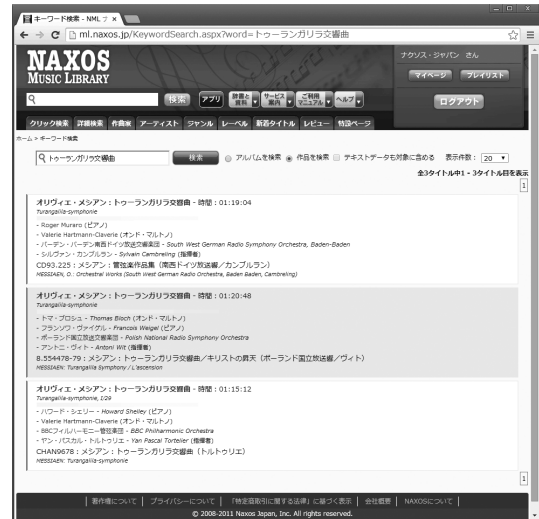


図2 トラックグループでの検索結果

本稿執筆時点で、このような検索機能はいまだ一般的ではない。類似サービスをすべて調査したわけではないが、たとえば我が国の国立国会図書館近代デジタルライブラリーの「歴史的音源」コーナー*7も、またロンドン大学キングズ・カレッジの音源アーカイブCHARM*8も、検索結果を楽章単位のリストで与えている。一方、大英図書館のサウンド・アーカイブ*9は作品単位での検索と再生の機能を有する。ここでは楽章単位のファイルの区切りをもち、作品全体を結合したファイルが提供されている。CD再生時のようにトラックをスキップすることはできないが、再生画面には波形データが表示されるので、ちょうどアナログ盤で盤面の溝の様子によって楽章の切れ目を視認するように、楽章や場面の切れ目を視覚的に探し、ランダムアクセスすることができる。

3.3 作品と録音

上記の考察から分かるように、クラシック音楽作品の資料を扱う場合、「作品」と「録音」の一対多関係は大きな意味をもつ。音楽愛好家や演奏史・録音史の研究者でなければ、多くの場合、利用者がまず探すのは個別資料としての「録音」ではなく「作品」である。この傾向理解は、NML日本語版のインターフェイス設計の大きな要因になっている。NML日本語版では、作曲家ごとに設けられたインデッ

*6 曲名の由来となったサンスクリット語に從うならば「トゥランガリラ」*7 <http://rekiion.dl.ndl.go.jp/>
*8 <http://www.charm.rhul.ac.uk/>
*9 <http://sounds.bl.uk/>

クスページに、その作曲家の作品が一覧で表示される。作品名を選ぶと、その作品の録音がリストされ、利用者はそこから任意の音源を選ぶようナビゲートされる。このインターフェイスはクラシック音楽に馴染みのない利用者に好まれると同時に、音楽愛好家には「聴き比べ」を容易にする機能として親しまれている。なお、大英図書館のサウンド・ライブラリーも、同様のインターフェイスを提供している。

このインターフェイスは、「作品」と「録音」の関係が書誌データベースで適切に管理されていなければ実現できない。この関係は、国際図書館連盟が「書誌レコードの機能要件」(Functional Requirements for Bibliographic Records, FRBR)で提示するデータモデルにおける、「著作」(work)と「表現形」(manifestation)の関係に相当する^{*10}。

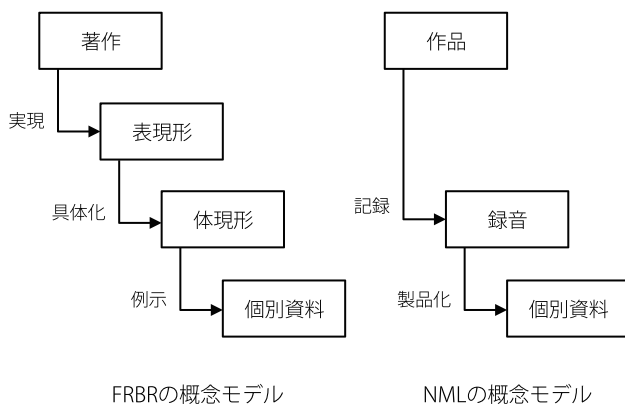


図3 作品と録音の関係

NMLの書誌データベースもまたこれに類する思想に基づいて設計されている。「著作」および「表現体」に相当する「作品」、「体现形」に相当する「録音」、そして個別資料は一对多対多の関係で管理されている。

4. 検索システム

4.1 ディスカバリー・サービスの台頭

近年、国内の図書館においても、従来型 OPAC から「ディスカバリー・サービス」と呼ばれる次世代型システムへの移行が次々に行われている。ディスカバリー・サービスを特徴付けるのは、(1) 外部リポジトリの横断検索、(2) 物理的形態が異なる資料の統合検索、(3) 検索ワードのトークナイズや同義語処理、検索結果の属性(ファセット)による絞り込み、関連資料レコメンド等の機能に代表される、情報発見を支援する機構である^{*11}。研究機関でこの種の検索サービスが普及すると、学術情報の検索サービスに対す

^{*10} Cf. 文献 [4], pp. 12–13. 訳語は文献 [5] による。

^{*11} ディスカバリー・サービスの定義や特徴付けについてはさまざまな見解が提出されている。ここでの特徴の列挙は筆者によるものであるが、ディスカバリー・サービスを定義しようとするものではない。その他、文献 [6][7] 等参照。

る利用者の期待も変化すると予想される。つまり、ディスカバリー・サービスの便利な機能に慣れ親しむようになると、その他の学術サービスにもそれを求めるようになると考えられる。それは一時的な変化かもしれない。しかし、利用者の要求が変化するならば、サービス提供者もそれに合わせるべきであろう。NML はリリース時より、従来型 OPAC の検索画面を模した「詳細検索」機能を搭載している。しかし従来型 OPAC が今後「時代遅れ」になるのであれば、NML もそれに合わせて次世代の検索サービスを提供しなければならない。かつて Google の圧倒的な利便性が図書館の資料検索システムにとって脅威であったが、いまや、NML にとってはディスカバリー・サービスの方が脅威である。

筆者は 2013 年に、試みに NML のすべての書誌データを、図書館の一般的な書誌データ記述形式である MARC21 フォーマットに書き出し、オープンソースのディスカバリー・システムとして普及している VuFind^{*12} にロードしてみた。結果として、現在の NML より使い勝手がよく、しかも高速で動作するウェブサイトが簡単にできてしまった。しかし同時に、多くの問題点に気付くことにもなる。とくに大きな問題は、MARC21 の思想と前節で述べた NML の思想とが根本的に異なるという点にあった。MARC21 は、図書館の所蔵する資料のメタデータを記述するためのフォーマットである。したがって、NML でいうところの「アルバム」が最終的なデータの単位になる。そこに含まれる「トラックグループ」や「トラック」は内容細目 (Contents Note) として記述できるが、検索結果はどうしても「アルバム」単位のリストになる。少し工夫をすれば、「トラックグループ」を資料の単位とし、「アルバム」を「シリーズ」に相当するデータとして保持することも可能であるが、データ設計の思想が本質的に異なる以上、いつか破綻を来すのは明らかである。NML のデータ形式を無理に MARC21 に合わせて、前節で述べた利点を放棄することは利点が少ない。MARC ベースの音楽資料 OPAC では実現が困難な要求を利用者は多数抱えていることは、文献 [8] のような実証的研究でも示されている。

4.2 Solr によるディスカバリー支援機能の実装

NML 日本語版では既成のディスカバリー・システムを利用することは断念し、VuFind のバックエンドを担うオープンソースの全文検索エンジン Apache Solr^{*13} のみを単独で採用することにした。Solr を若干カスタマイズして利用すれば、本節冒頭で掲げた情報発見支援機構は実現できる。とくに、前掲 3 種の特徴のうち、NML にとって最も優先度が高い 3 番目の諸機能は Solr の標準機能ですべて実

^{*12} <http://vufind.org/>

^{*13} <https://lucene.apache.org/solr/> Hathitrust が採用していることでも知られる。

装できる。本稿執筆時点ではまだ試作段階であるが、本稿公表までに、NML 日本語版では検索システムを Solr を利用した新システムでリプレースする計画である。

NML 日本語版へのディスカバリー支援機能の導入目的は、単に図書館システムの動向に追随することのみにあるのではない。それらの機能が、学術データベースの利用者の利用目的に適うものであると判断されるからでもある。NML の旧検索システムを含め、多くの従来型の検索システムでは、資料の属性による検索結果の絞り込みができず、検索ワードを追加して再検索することにより結果を限定していくというインターフェイスが取られている。これは、探している資料の書誌情報が分かっており、その資料をピンポイントで探すときには有効であろう。しかし、書誌情報が分からず、或いはそもそもどのような資料が世の中に存在するのかも分からずに、漠然と「ワーグナーのオペラってどんなものか聴いてみたい」というような要求を満たすにはふさわしくない。「ワーグナー」で検索しても、さらに「オペラ」というワードを追加しても、まだまだ検索結果は多すぎる。そこからさらに、レーベルや演奏者について、「なんとなく知っている名前を選ぶ」という仕方でもよいから絞り込み、自分の目的に合いそうな資料を探していく。NML 日本語版は、このような情報発見のプロセスを支援しようとしている。研究者ならば書誌情報くらい調べてから図書館へ来い——そのような批判もあるだろうし、その研究者の専門領域に関してはその批判は或る程度は妥当であろう。しかし NML は、音楽の専門家でない研究者全般に対して、音楽資料を研究に取り入れてもらいたいと考えている。ワーグナーがインド哲学の影響を受けていることをどこかで聞いてきたインド学研究者が、試しにワーグナーを聴いてみて、また NML で提供されているブックレット（ライナーノーツ）も読んでみて、研究に新たな視点を取り入れる。このようにして、サービス運営者側も、利用者をディスカバーすることができる。

もっとも、NML 日本語版の検索システムはディスカバリー機能に特化するのではなく、これまでどおり、ピンポイントで特定の資料を取得したい専門的な利用者の要求にも応えていく。従来型検索とディスカバリー指向の検索は、常に共存しつつあるだろう。NML の設計においては、それぞれを「既知の録音の検索」と「未知の録音の検索」と規定している。前者においては捜している資料を短時間で的確に提示することが目標であり、そのため、検索ワードの形態素解析といった過剰な検索支援は無効にする。後者においては、情報発見の可能性を最大限に高めることが目標である。未だ知らない作品や録音を求めて検索するとき、最終的な音源資料だけでなく、検索結果の画面すべてが利用者にとって情報となる。そこに含まれるノイズを減らし、かつ潜在的な利用価値をもつ資料をなるべく多く提示できるよう、いまだユーザー辞書のチューニングを続け

ている。

5. まとめ

以上で議論してきたことを、普遍的な視野で整理しよう。デジタル・アーカイブやデジタル・ライブラリーにおいて、サービスの利用を拡大し、諸科学の発展に貢献しようとするならば、情報検索のインターフェイス設計はきわめて重要な仕事となる。そしてその設計においては、資料の物理的形態や、コンテンツ制作のプロセスに由来する先入観を排除するべきである。本稿では、音楽コンテンツ配信サービス Naxos Music Library が音楽愛好家の趣味的サービスに留まらず研究者の研究活動に寄与できるデジタル・ライブラリーとなるため、「CD 音源のリポジトリなのだから検索結果は CD のリスト」「CD はトラックで構成される」といった常識を見直し、「利用者が何を求めればその要求が満たされるのか」を考え直す試みを提示した。利用者は情報を求めている。そして、情報は、検索の最終地点であるコンテンツデータのみにあるのではない。検索結果のひとつひとつが情報であり、その情報にもとづいて、次のアクションを定め、最終的にコンテンツデータに至る。あるいはコンテンツデータに至らずとも、そのようなコンテンツは存在しないという知識をもたらず空の検索結果が得られれば、それも情報である。本稿では、Naxos Music Library において、物質的実体を伴わない「トラックグループ」というデータ単位を設け、また「作品」と「録音」の関係を記述し、豊かな情報を生み出し得る仕組みを用意したことを紹介した。他のデジタル・アーカイブやデジタル・ライブラリーにおいても、制作者側が情報の基本的単位と見なしている実体が、ほんとうに利用者にとって意味のあるものなのかを問い直してみる価値があるだろう。

参考文献

- [1] 及川昭文. 「特定領域研究『人文科学とコンピュータ』からの報告」 <<http://www.classics.jp/RCS/NL03/NL03Oika.pdf>> (Accessed on April 27, 2014).
- [2] 永崎研宣. 「デジタルアーカイブの弁証法」『社団法人 情報処理学会 研究報告』2005-CH-68. pp. 17-24. 2005.
- [3] Ana Dubnjakovic and Patrick Tomlin. *A Practical Guide to Electronic Resources in the Humanities*. 1st edition. Cambridge: Chandos Publishing. 2010.
- [4] IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records*. München: K.G. Saur. 1998. <<http://www.ifla.org/files/assets/cataloguing/frbr/frbr.pdf>> (Accessed on April 27, 2014).
- [5] IFLA Study Group on the Functional Requirements for Bibliographic Records 著. 和中幹雄・古川肇・永田治樹訳. 『書誌レコードの機能要件 IFLA 書誌レコード機能要件研究グループ最終報告 (IFLA 目録部会常任委員会承認)』日本図書館協会. 2004.
- [6] 片岡真他. 「図書館の検索インターフェースとユーザ支援技術」『メディア教育研究』第7巻第2号. S19-S31. 2011.
- [7] Tamar Sadeh. “From Search to Discovery”. In the on-

line proceedings of IFLA World Library and Information Congress 2013. <<http://library.ifa.org/104/1/098-sadeh-en.pdf>> (Accessed on April 27, 2014).

- [8] 金井喜一郎. 「音楽資料に関する OPAC 検索機能要件—レファレンス記録の分析を通じて—」『日本図書館情報学会誌』 56(4). 2010.

【正誤表】

2014-CH-102 NO. 7

1 ページ目

英語タイトル

誤 : Services

正 : Service