

利用者の音楽情報要求からみたメタデータ要素の有用性

—FRBR, FRAD, Variations を対象に—

金井喜一郎（慶應義塾大学大学院 kiichirokanai@a3.keio.jp）

1. 研究の背景と目的

音楽資料に対するメタデータの開発・研究が盛んに行われている。メタデータは本体に対する代理物 (surrogate) であるから、メタデータ要素の数を増やせば、それだけ音楽資料のより詳細な記述が得られることになる。しかしながら、利用者の実際の情報要求に照らして、膨大な数のメタデータ要素が本当に必要かどうかは自明ではない。実際の要求に照らして、その要素の必要性を検討することは、システムの実装において重要であり (図 1)、これを明らかにするには、利用者の実際の情報要求を何らかの方法で集め、それらを満たすのに必要なメタデータ要素を特定することが考えられる。この線に沿った研究としては、Lee(2010)¹⁾ および金井(2010)²⁾がある。

本研究では、金井(2010)によりも、さらに洗練された結果を得るために、次の3点でその方

法を拡張する。①メタデータモデルの調査対象を広げ、金井(2010)で取り上げたFRBRとVariations(後述)のほか、新たにFRADを追加、②利用者の情報要求に関するデータの安定性を計るため、新たに2つの図書館のレファレンス記録を追加(合計6つの記録)、③より詳細な分析のために、メタデータ要素がレファレンス記録中出现する頻度を集計する。

2. 先行研究

2. 1 Lee(2010)による利用者の音楽情報要求の分析

Lee(2010)は、現在の音楽情報検索研究における大きな問題の1つが、実際の利用者とその情報要求に関する実証的な調査の不足であるとして、Googleが運営するQ&AサイトであるGoogle・アンサーズから実際の利用者の音楽に関するクエリのログを抽出し、そこに含ま

れる情報要求の素性を分析した。その結果、人名、タイトル、日付、ジャンル、役割、歌詞、場所など少数の素性への集中が見られた。

2. 2 金井(2009)³⁾による音楽分野のレファレンス記録の分析

金井(2009)は、音楽分野のレファレンス記録(NDLレファレンス協同データベースおよび昭和音楽大学附属図書館の2つの記録)から、楽譜や音楽録音資料などに特有の情報要求(検索要求)を分析の上、それらを25の検索課題として整理

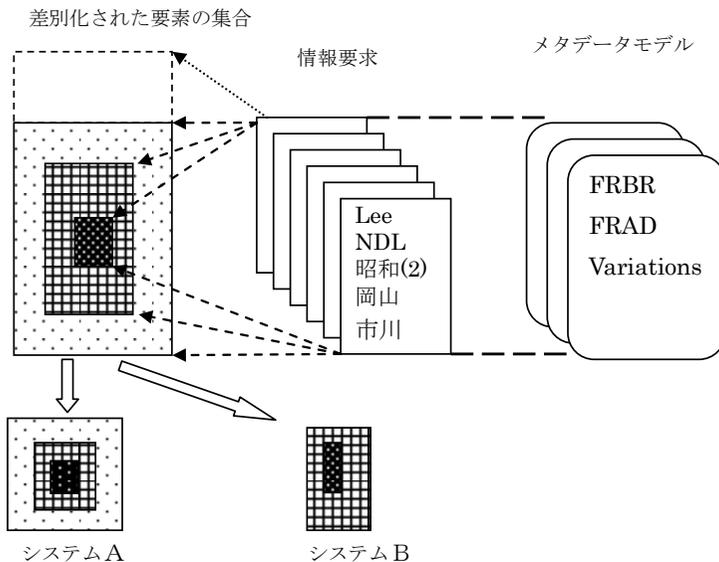


図 1 情報要求とメタデータ要素の関係

し、これらの検索課題を用いて、実際の音楽図書館の OPAC の検索機能を調査した。そして最終的に、調査結果に基づいて、音楽資料（特に楽譜や音楽録音資料）に関する OPAC 検索機能要件を導出した。

2. 3 金井(2010)によるメタデータ要素の有用性の分析

前述した Lee(2010)および金井(2009)の研究成果を基とし、既存のメタデータモデルである FRBR, Variations の要素の有用性（どれだけ使用されるのか）について、実際の利用者の情報要求から調査した。その結果、Lee(2010)と同様に、一部の要素が集中的に使用されていることが分かった。

しかし、「今回の調査では各特徴の出現回数を考慮していないため、この区分は一つの目安に過ぎず、正確な区分とは言いがたい」²⁾のものであった。

3. FRBR, FRAD, Variations プロジェクト

3. 1 FRBR おび FRAD

FRBR (Functional Requirements for Bibliographic Records)⁴⁾は、1998年に国際図書館連盟(IFLA)により勧告された概念モデルで、実体(著作, 表現形, 体現形, 個別資料, 個人, 家族, 団体, 概念, 物, 出来事, 場所)と実体間の関連によって、そのデータモデルが示されている。FRAD (Functional Requirements for Authority Data)⁵⁾は、FRBRを拡張するモデルであり、典拠データの機能要件が示されている。

3. 2 Variations プロジェクト

Variations は、「世界で最初のデジタル音楽図書館システムの一つ」⁶⁾である 1996年に始まったインディアナ大学のデジタル音楽図書館プロジェクトであり、Variations2 (2000-2005), Variations3 (2005-2008), Variations/FRBR (2008-2011 予定)へと引き継がれていく。ここでは、システムの開発とともに音楽に特化し

たメタデータモデルの開発が進められた⁶⁾。最新版の Variations/FRBR は FRBR および FRAD の要素を取捨選択し、そこに独自の要素を加えたもので、本研究では、このモデル⁷⁾を基にし、そこに必要に応じて Variations2 および Variations3 の記述メタデータ要素⁸⁾⁹⁾¹⁰⁾¹¹⁾を加え、それを調査対象とした。

4. 調査の枠組みと方法

4. 1 調査の目的, 対象

金井 (2010) では、Lee (2010) が分析した利用者の情報要求の素性 (タイトル, アーティスト名, ジャンルなど) の一覧に、金井 (2009) の 25 の検索課題を加え、これを利用者の音楽情報要求を表すものとして調査に使用した。しかし、後者に関しては、レファレンス記録を素性に分解 (整理) する過程で余計な作業が加わっているため、レファレンス記録に表れている利用者の情報要求が縮小されてしまう恐れがある。よって、より正確に情報要求を捉えるため、本研究では、金井 (2009) については、原データであるレファレンス記録を分析する。

また、利用者の音楽情報要求に関するデータの安定性を計るため、新たに 2 つのレファレンス記録を分析の対象として金井 (2009) に加えた。2011年1月15日の時点でインターネット上で閲覧が可能な記録のなかから、金井 (2009) で使用した記録と重複せず、さらに NDC760 で音楽分野の抽出が可能な記録を選ぶこととした。その結果、岡山県立図書館のレファレンスデータベースおよび市川市立図書館のレファレンス事例集が調査の対象に加わった。

4. 2 音楽情報要求のメタデータ要素への対応づけ

本研究では、レファレンス記録に含まれる情報要求を満たすために必要なメタデータ要素を判別し、その数を集計する。これによって、メタデータ要素が情報要求を満たすためにどれだけ実際に使用されているかを明らかにすることができる。同時に、情報要求を満たすために必

要でありながらも、既存のメタデータモデルには含まれない要素を特定することも可能である。一方、Lee(2010)は、同様な作業を、グーグル・アンサーズの記録に対して行っている。しかしながら、その作業は十分に細かくないため、今回改めて、その作業を繰り返した。ただし、本研究では、Lee(2010)が使用した記録そのものには立ち返らず、Leeにより抽出された素性の一覧を利用して、各メタデータモデルの要素の使用の程度を算出した。金井(2010)では、対応づけられた素性の数を単純集計したのみだが、本研究では、利用者の情報要求の実際の記録での出現頻度をも集計する。これは、利用者の情報要求を、より正確に把握するためである。

5. 結果と考察

ここでは、記録中に出現すればするほど、そのメタデータ要素は利用者にとって有用であると仮定し、その回数を4段階に集計する。具体的には、各要素に対応付けられたレファレンス記録を計数し、Lee(2010)を参考に、25%以上を「有用性が高い：■」、10%以上25%未満を「有用性が中位：□」、1%以上10%未満を「有用性が低い：○」、1%未満を「有用性なし：無印」とした。さらに○の件数が多いため、便宜上、○を2段階(4%以上10%未満：○+, 1%以上4%未満：○)に区分した。なお、「有用性なし：無印」を1%未満に設定したのは、Lee(2010)の記録では、1%未満のデータが示されていないためである。

表1にその結果を示す。表1は、6つの記録すべてを合計し、その値に基づいて、各データ要素が全出現に占める割合を示したものである。例えば、最も出現する割合が高いのは、「著作のタイトル、表現型のタイトル」であり、全記録の66.5%を占めていた。表1では、出現する記録の割合が1%以上ものを表示しており、その要素数は41件である。この41件は利用の情報要求を満たすために重要な要素であると考えられる。さらに、その区分に着目すると、■：5件、□：4件、○+：9件、○：23件であり、

特に■の要素はコアとなる要素であるといえる。

次に、これらの41件のデータ要素がどれだけ含まれているかという点から、各データモデルの評価を試みる(表2)。評価にあたっては、各記号を点数化し、その総得点を比較した。その結果、FRBRとFRADがほぼ同じ点数である一方、Variationsはそれらの約1.4倍の点数であり、全41件中39件を含んでいることが分かった。

表2 各データモデルの点数比較

区分(点数)	全体		FRBR		FRAD		Variations	
	件数	点数	件数	点数	件数	点数	件数	点数
■(3点)	5	15	4.5	13.5	3	9	5	15
□(2点)	4	8	2	4	3	6	4	8
○+(1.2点)	9	10.8	6.5	7.8	6.5	7.8	8	9.6
○(0.9点)	23	20.7	14.5	13.1	17	15.3	22	19.8
合計	41	54.5	27.5	38.4	29.5	38.1	39	52.4

* ○については、平均で1点となるようにした上で、○+と○の点数に差をつけた。

最後に、表1のデータの安定性について確認する。つまり、この方法の場合、新たなレファレンス記録が追加されれば、当然、メタデータ要素の順位や割合は変動することが予想される。この変動は必然ではあるが、どれだけ表1の結果が安定しているかどうかを見るために、LeeとNDLから出発し、順次、データを追加してみた場合に、対応する記録が1件以上存在する要素がどのように純増するかを集計した(表3)。全体では、3つの記録の合計(表3の②)以降は、ほぼ変化がなく、□も同様である。■については、2つ記録の合計(表3の②)以降、全く変化がない。○(○+含む)も4つの記録の合計以降(表3の③)は、合計数に変化はない。以上のことから、データの安定性は高いと考えられる。

表3 記録数の増加に伴うメタデータ要素数の変化

No.	項目	出現する記録が1件以上	メタデータ要素数			
			■	□	○+	○
①	Lee+NDL	69	5	4	15	25
②	①+昭和レファレンス	72	5	3	13	22
③	②+昭和OPAC	72	5	3	11	21
④	③+岡山	72	5	3	11	21
⑤	④+市川	73	5	4	9	23

6. 結論

利用者の音楽情報要求を既存メタデータの要素へ対応づけることにより、有用性によるメタデータ要素の差別化を試みた結果、大きく4段階に区分することができた。今回は、金井(2010)

表1 出現する記録の割合が1%以上のメタデータ要素

No.	要素	割合(%)	有用性
1	著作のタイトル、表現形のタイトル	66.5	■
2	キャリアの形態	49.6	■
3	責任表示、個人名	47.7	■
4	(関連)創造(個人と著作)	45.6	■
5	著作の形式(形式やジャンルを含む)、表現形のジャンル/形式/スタイル	42.0	■
6	著作の主題(⑤~⑦の合計)	19.9	□
7	演奏手段(音楽作品、楽譜または録音)(④含む)	15.7	□
8	(関連)実現(個人と表現形)	11.4	□
9	著作の主題(音楽関係印刷資料の主題)(⑦)	10.7	□
10	個人名	6.8	○+
11	番号表示(音楽作品)	6.1	○+
12	楽譜の種類(楽譜)	5.7	○+
13	著作の言語、表現形の言語(⑨~⑩含む)	5.2	○+
14	著作の主題(音楽作品の主題)(⑤)	4.9	○+
15	出版者・頒布者	4.5	○+
16	表現形のタイトル	4.5	○+
17	著作の主題(音楽作品の解説)(⑥)	4.4	○+
18	(関連)翻訳	4.0	○+
19	(関連)全体と部分(著作、表現形)(⑬含む)	3.9	○
20	著作の日付(①~③の合計)	3.9	○
21	個人の伝記/経歴	3.7	○
22	表現形の言語(⑩)	3.4	○
23	個人の性別	3.3	○
24	著作に関する場所(発祥地、作曲の場所、初演の場所)(⑧含む)	2.8	○
25	出版日付・頒布日付	2.5	○
26	演奏手段(楽譜または録音)(④)	2.5	○
27	個人の職業	2.4	○
28	表現形のその他の特性[同一著作の異なる表現形を識別するのに役立つ特性](editionやversion含む)	2.4	○
29	(関連)追補(著作、表現形)	2.2	○
30	著作の発祥地(⑧)	2.1	○
31	表現形の成立日付	2.0	○
32	(関連)編曲	1.9	○
33	個人に関する場所(出生地、死亡地、国、居住地)(⑪~⑫含む)	1.8	○
34	個人の日付	1.4	○
35	著作のその他の特性[同一タイトルをもつ別の著作を識別するのに役立つ特性](インデックス含む)	1.4	○
36	団体名	1.4	○
37	調(音楽作品)[その著作が最初に作曲されたときの調、表現形の調]	1.3	○
38	著作の来歴	1.2	○
39	内容の要約(表現形)	1.1	○
40	著作の言語(⑨)	1.0	○
41	(関連)改作(著作、表現形)	1.0	○

では考慮しなかった要素の出現頻度をも考慮し、より詳細な分析ができた。

さらに、分析対象のレファレンス記録を増やすことで、本研究で整理した利用者の情報要求データの安定性が高いことも分かった。

今回の調査の過程で、利用者の音楽情報要求の基とした各記録に、ジャンルごとの特徴が表れていることが分かった。これらを詳細に分析し、館種や音楽ジャンルごとにメタデータ要素の有用性を探ることが、今後の課題である。

注・引用文献

1) Lee, Jin Ha. Analysis of user needs and information features in natural language queries seeking music information. Journal of the American Society for Information Science and Technology. 2010, vol. 61, no. 5, p. 1025-1045.

- 2) 金井喜一郎. 利用者の音楽情報要求からみたメタデータ要素の有用性: FRBR, Variations を対象に. 三田図書館・情報学会研究大会発表論文集. 2010, vol. 2010, p. 53-56.
- 3) 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. 三田図書館・情報学会研究大会発表論文集. 2009, vol. 2009, p. 73-76.
- 4) IFLA Study Group on Functional Requirements for Bibliographic Records. Functional Requirements for Bibliographic Records. Final Report. K. G. Saur, 1998, 136p. <http://www.ifla.org/files/cataloguing/frbr/frbr.pdf>. (accessed 2010-9-5).
- 5) Patton, Glenn E; IFLA working group on Functional Requirements and numbering of Authority Records (FRANAR.). Functional Requirements for Authority Data: a Conceptual Model. München, K.G. Saur, 2009, 101 p., (IFLA Series on Bibliographic Control, 34).
- 6) Variations3: an integrated digital library and learning system for the music community. p. 2. http://www.dlib.indiana.edu/projects/variations3/docs/Indiana_University_IMLS_2005-02-01.pdf, (accessed 2010-9-5).
- 7) XML schema definitions for FRBR, version 1.0. <http://www.dlib.indiana.edu/projects/vfbr/schemas/1.0/index.shtml>, (accessed 2010-9-5).
- 8) IU digital music library data model specification V2: final release. 2003. <http://variations2.indiana.edu/pdf/DML-DataModel-V2.pdf>, (accessed 2010-9-5).
- 9) Davidson, Mary Wallace; Hemmasi, Harriette; Minibayev, Natalia. Indiana University Variations2 digital music library project: controlled vocabularies. 2002. <http://variations2.indiana.edu/pdf/DML-vocab-notes.pdf>, (accessed 2010-9-5).
- 10) Riley, Jenn; Hunter, Caitlin; Colvard, Chris; Berry, Alex. Definition of a FRBR-based metadata model for the Indiana University Variations3 project. 2007. <http://www.dlib.indiana.edu/projects/variations3/docs/v3FRBRreport.pdf>, (accessed 2010-9-5).
- 11) Riley, Jenn; Mullin, Casey; Colvard, Chris; Berry, Alex. Definition of a FRBR-based metadata model for the Indiana University Variations3 project phase 2: FRBR group 2&3 entities and FRAD. 2008. <http://www.dlib.indiana.edu/projects/variations3/docs/v3FRBRreportPhase2.pdf>, (accessed 2010-9-5).