情報メディア組織化への天の邪鬼的アプローチ

谷口祥一
図書館情報学系助教授

情報メディアの組織化とは

われわれが何らかの情報を包含していると考えるリソース（資源）が「情報メディア」であり，それら多種多様なものをいかに多様な要求に効率的・効果的に応えられるように秩序づけるのか（特には，事前に秩序づけておくのか）というのが「組織化」である。ここには，情報メディア自体の組織化（加工，編集等）と，情報メディアから独立した記録を作成し，それを組織化するという2つの方法がある。前者は，情報メディアが電子メディアになって初めて，ある程度の処理の自由度が得られており，それ以外の媒体固定のメディア（媒体と不可分のもの）については，媒体としての保管・保存，あるいは排列等による提供などに処理の範囲は限定される。

後者の方法，すなわち情報メディアから独立したサロゲート（書誌データ，目録データ，メタデータなどと呼ばれる）を個々の情報メディアについて作成し，それを組織化し提供することによって，情報メディアの発見・効率的なアクセス，効果的な利用，あるいは長期の保存も可能にしようとするのが，私の依拠する基本的な枠組みでもありフィールドでもある。こうしたことはあらゆる領域，組織・機関において広く行われてきており，その点では特に目新しさはない。例えば，図書館，文書館，博物館・美術館，さらには情報センターやデータセンターといった組織では，かなりの労力と経費をかけてこのような組織化が実現されている。

現在は，特にインターネットに代表されるネットワーク上のリソースの発見・活用に関わりメタデータというものが注目されており，さらにはそれを用いたsemantic webなるものが関心を集めている。これも先の区分からいえば，同一方向への展開といえよう。

私の信条？

前述のような基本的な枠組みに依拠した上で，具体的な研究テーマの設定とその問題解決の試みにおいて，私はどのような攻め方を取っているのか。一言でいえば，それは流行（はやり）や主流からは距離を置くというものである。流行や主流に首を突っ込んだ時点で，自らの鈍い頭を使って自分のペースでものを考えることがやりにくくなるのは，私の場合明々白々であり，それでは領域への貢献も望むべくもない。それゆえ，先の攻め方が唯一可能な選択肢ともなるが，同時に生来の天の邪鬼な気質に具合よく合っているようである。あとは，自らの直感や感覚に頼って取り組むべき問題を設定し，使用可能な手段や道具と引き比べながら解決法をあれこれと模索するだけである。そして最終的には領域に貢献するというのが，私の信条である（今のところ気概にすぎず，十分に貢献できておりませんが。）

これまでのところ，図書館情報センター等で作成される「目録」なるものを最初の取りかかりにして，それが抱える問題に対して先の攻め方を取ってきた。その顛末の一部を以下に示そう。

対象事象・領域のモデリング

サロゲートを設計し作成するには，形式化して考えるならば，まず組織化対象の情報メディアが形成する事象や領域をどのように捉えるのか，モデルを構築し検討することが必要となる。当然のことながら，モデルは組織化の目的や目標あるいは制約条件に拘束される。これら要求定義は，特定の環境に即したときには具体的かつ明確な条件として提示されることになろうが，領域全体を視野に入れた議論のレベルでは抽象的なものとならざるをえない。加えて，情報メディアのライフサイクルの各段階ごとに多少とも異なる要求として現れる。

要求定義を踏まえた対象事象のモデリングは，まずは概念設計として行われ，その結果が以降のデータ項目やレコードの設計（場合によっては，データベースのスキーマ設計）を拘束することになる。従って，この概念レベルのモデリングがまずもって重要となる。それぞれの領域においてこのような概念モデルの構築や検討が現在精力的に行われており，図書館の目録については，国際図書館連盟（IFLA）によるFRBR（Functional Requirements for Bibliographic Records）モデルが最も完成度が高いものと評価されている。事実，同モデルを基盤にし

て，直面する具体的な問題群に対する解決への議論や取り組みが行われている。

私自身は，こうした概念モデリングが隆盛を見せる以前から，同様の意図の下に単純なモデルを構築し提案してきた。当時は興味を示してくれる人も少なく（今でも同じ状況？），またFRBRのような完成度にまで自力で展開し洗練することもできなかった。自らの非力を悔やむと同時に，基本的な方向性において間違っていなかった点を確信できたことを救いと考えている。今でも，FRBRの草案を初めて目にしたときの驚きを鮮明に覚えている。

そこで，自ら示した単純なモデルの提案に含まれていた主張の一つが，テキスト（テクスト：text, expressionなどと呼ばれる）レベルの重視という点にあったことを踏まえ，FRBRが公表された後は，この点のみの提案と議論に焦点を絞る戦略に切り替えることにした。テキストとは，情報メディアの著作者が文字・数字・記号等を用いてその知的内容を表現したレベルを指し，それがフォーマットやレイアウトを伴って媒体に記録されたレベルとは区別する。このテキストレベルの重視とは，同一テキストが多様なフォーマットや媒体で出現する事態への対応，テキスト自体に関わる豊富な情報提供，さらには電子メディアやそのメタデータの議論との親和性を意味し，現在でも有効な考え方と個人的には考えている。その点にこだわり，このような考え方を採用していないFRBRに対抗したモデルの提案（実質的にはFRBRへの修正提案），さらには提案モデルから導かれるデータ作成法の検討や実験（新規作成や既存データからの変換可能性の検証）を現在大急ぎで進めている。これは従来の枠組みに対する大きな変更を意味することもあり，現在のところ同調者は皆無に近い。

処理プロセスのモデリング

サロゲートが対象とする事象・領域のモデリングと併行して，サロゲートの作成処理プロセス（処理タスク）のモデリングを現在進めている。先の対象事象に対するモデリングの結果，個々のデータ項目が設定された段階では，どのようにしてその値を記録するのかが問題となる。最も簡便には，値の採取・記録法を一切決めず，すべてサロゲート作成者の判断に委ねることになろうが，これでは高品質かつ多様な用途に使用しうるサロゲートは期待できない。それゆえ，目録の場合にはデータ項目ごとに値の採取法，記録法をルールの形式で詳細に定め

ている。

こうした作成処理プロセスさらにはそれに対応するルールセットをいかに設計するのかという課題に対して，私は先の対象事象のモデリングと同様に，ここでも概念レベルのモデリングを行い，それに基づき詳細設計に展開させることを試みている。このような試みは，喜ぶべきか，私の他にはこれまでのところ皆無である。

その基本的な考え方とは，ルールセットの概念設計を，要求定義，基本モデル（コアモデル）構築，基本モデル展開，展開モデル限定化の各フェーズに分けて実現しようとするものである。処理プロセスの基本的かつ汎用的な概念モデルである基本モデルの構築を基盤に据え，それを個別データ項目の特性に合わせて展開することによって，さらには私が以前に提案した「指向性（orientedness）」という概念装置を要求定義から特定のシステムや環境に適合したルールセット獲得（展開モデルの限定化）まで媒介させることによって，汎用性と一貫性，そしてスケーラビリティを備えた設計方式とすることができると考えている。

おわりに

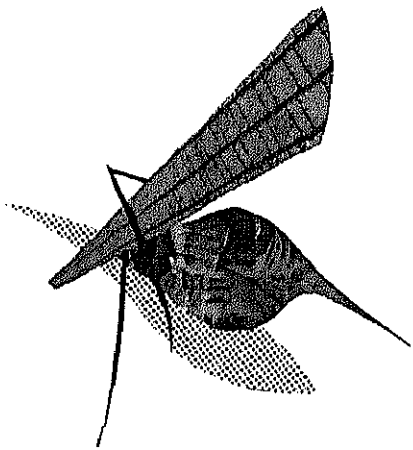
極めて偶然的な恩師との接点から大学教員の道を歩むことになった私であるが，これまでに幾ばくかでも成果を出すことができたとすれば，それは選択したテーマにこだわり続け，そして主流に飲み込まれず，異なる解決法を求めてきた結果といえよう。徒手空拳で出発した私にとって，それは唯一可能な選択肢であった。

対象事象のモデリング，処理プロセスのモデリングとも突き詰めていくと，知識のモデリングに行き着くように思われる。そして，それは現在流行のsemantic webなどでも注目されているオントロジーの問題と言い換えることもできよう。ここでもまた，主流に飲み込まれず，それらとは異なる接近法を見出すことが，あるいは見過ごされている側面を取り上げることが私にとって不可欠となる。これから先暫くは，天の邪鬼でいかざるをえないようである。

同時に，より汎用的な，多少とも射程距離の長い研究を目指したいとも考えている。これまでの成果は対象としてきた目録の範囲を多少とも越えて適用可能と考えてはいるが，何せその射程距離には限界がある。他の組織や領域においても成立する，より普遍的な課題に対して，より汎用的な解決法をもって答えるという構図に手を伸ばすことも考えている。

そのためのアイディアの種も準備しているが，なかなか時間的に手が出せないのが実状である。

大学統合後，筑波大学という総合大学の中で研究を展開し，他分野の人たちと互していくことに対して，よい意味での緊張感を感じている。どこまで通用するかは不明であるが。(個人のホームページに成果等は随時掲載していきたいと考えている。ご興味のある方は，そちらをどうぞ。

http://www.slis.tsukuba.ac.jp/~taniguch/)

(たにぐちしょういち　図書館情報学)