

(付 録)

海外電子図書館の実態調査報告

1. The University of Michigan に在外研究員として滞在して —— 63
研究開発室 今井助教授
2. ADL '99 に参加して ————— 68
研究開発室 今井助教授
3. ECDL '98 ————— 71
研究開発室 新助手
4. Dublin Core Workshop ————— 73
研究開発室 新助手

1. The University of Michigan に在外研究員として滞在して

本学附属電子図書館研究開発室で筆者（今井）は、1997年9月から1998年7月まで文部省在外研究員として米国 The University of Michigan に客員研究員として滞在し、電子図書館の研究を行った。ここでは、その報告を行う。

筆者は、School of Information に滞在した。実質受け入れ教官は Daniel Atkins 教授（当時学部長）であった。Atkins 教授は The University of Michigan において行われている電子図書館プロジェクト（University of Michigan Digital Library : UMDL）を推進する中心人物であった。UMDL は School of Information, University Library, Information Technology Division の大学内3部局が中心となって推進する電子図書館プロジェクトである。また同時に米国における国家プロジェクトとしての Digital Library Initiative (DLI)でも中心的な役割を果たしていた。（なお、DLI プロジェクトは NSF, 1998年に終了し、現在は DLI2 が行われている）

The University of Michigan での電子図書館プロジェクトは大きく2つのプロジェクトにわけることができる。一つは先ほどの DLI プロジェクトに参加しているもので、もう一つは School of Information, University Library そして Information Technology Division が中心となって推進しているプロジェクトである。どちらも UMDL というプロジェクトの名称であり、区別が付きにくい。実際、これら2つのプロジェクトは別個に行われているのではなく、Atkins 教授を中心として双方のプロジェクトに参加する研究者が定期的にミーティングを実施し、相互の意見交換や共通する題材で議論を行っていた。以下、それぞれのプロジェクトと関連するものについて調査研究を行ったのでこれを概観する。

DLI プロジェクトとしての UMDL

このプロジェクトでは、高等学校における「科学 (Science)」の授業において、インターネット上に開設されている Web page を有効に教材として利用することを直接の目的とし、それを実現するためのデジタルライブラリサーバ機構の研究を行っている。このプロジェクトは、NSF, NASA, DARPA が共同で設定した Digital Library Initiative の一つとして実施された。

米国では宇宙開発が大きな国家プロジェクトとして実施されており、火星探査 (Mar's Path Finder が有名) や土星や木星の探査 (Pioneer 計画が有名) など、数多くの探査が行われその成果が NASA などに貯えられている。米国の高等学校における科学の授業では、宇宙科学のカリキュラムが設定されており、これらのデータを利用した授業が実施されている。(NASA が DLI のスポンサになった動機の一つは、惑星探査により得られた成果を国民に還元することであるともいわれている) 授業に用いられる教材としての宇宙科学の研究成果は、全米各地のみならず世界中の関連する研究機関から Web により公開されている。授業ではこれらの分散している教材につ

いて、授業の進展度や生徒の興味に応じて取捨選択されて用いられる。この取捨選択も、教師が一方的に行うのではなく、必要な知識を与えた上で生徒に行わせる。効率的な授業を行うためには、分散している宇宙科学の研究成果を統一的に検索、選択するメカニズムが必要になる。UMDL ではこのメカニズムを情報科学、特に人工知能分野で研究が行われているエージェントを用いて実現し、合わせてエージェント技術の研究を行っている。図 1.1 にその概念図を示す。

Agent Architecture: 97

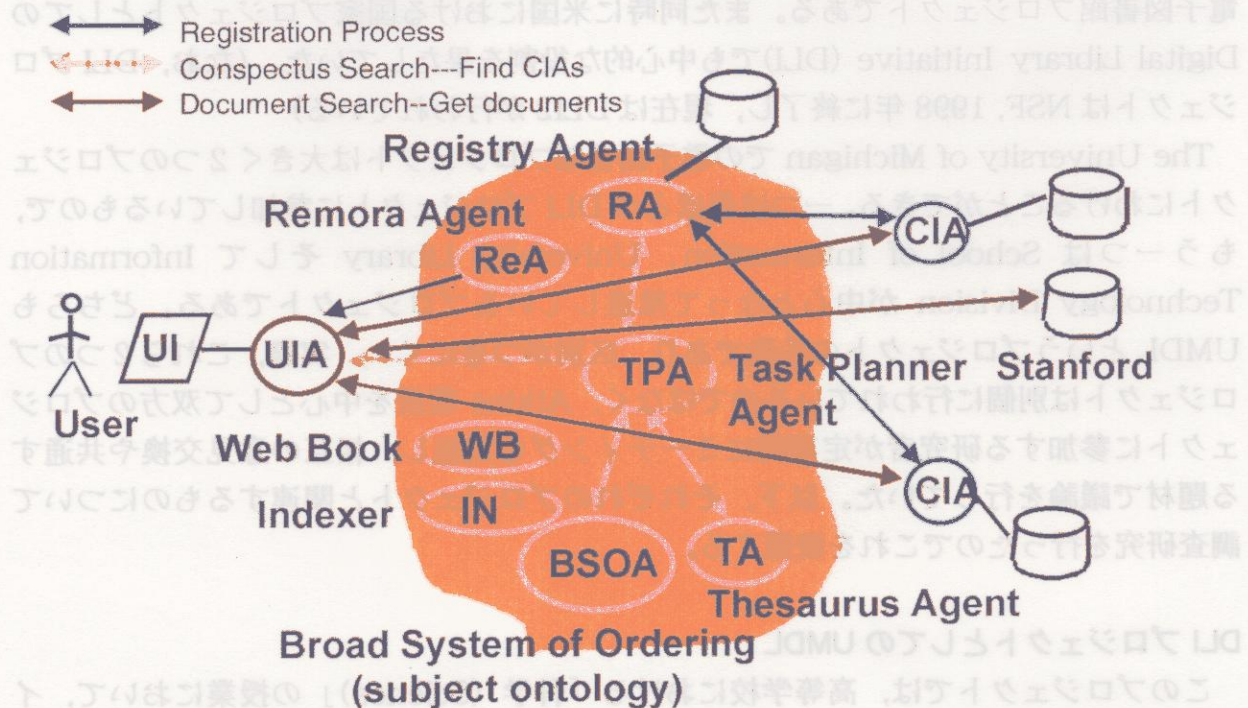


図 1.1 UMDL のエージェントアーキテクチャ (1997 年当時)

このエージェントアーキテクチャでは、エージェントは大きくわけて 3 種類のものに分割されている。User Interface を司るエージェント (UIA)、Mediator と呼ばれるエージェント、そしてサーバとの通信を行う Collection Interface エージェント (CIA) である。Mediator には図中にあるようにさまざまな機能を持ったエージェントが存在する。その中でも、Registry エージェントは CIA から各サーバの情報を収集し、保持し、それを他のエージェントからの要求に応じて伝える。その他の Mediator はそれぞれの目的をもったエージェントが存在し、互いに協調してユーザからの検索要求にこたえる。

図 1.2 にユーザインタフェースの例 (1997 年当時) のものを示す。これは <http://umdl.umich.edu/> から体験することができる。このシステムは実際に Ann

Arbor 地区にある2つの高校の授業で利用され、機能やインターフェースは授業での経験をもとに、研究者と高校教諭が話し合って決めている。授業の進行と時間の関係などで途中で中断しても再開できるように、途中結果の保存もできるようになっている。インターネットからデモバージョンを試用することができる。

(<http://www.si.umich.edu/UMDL>)

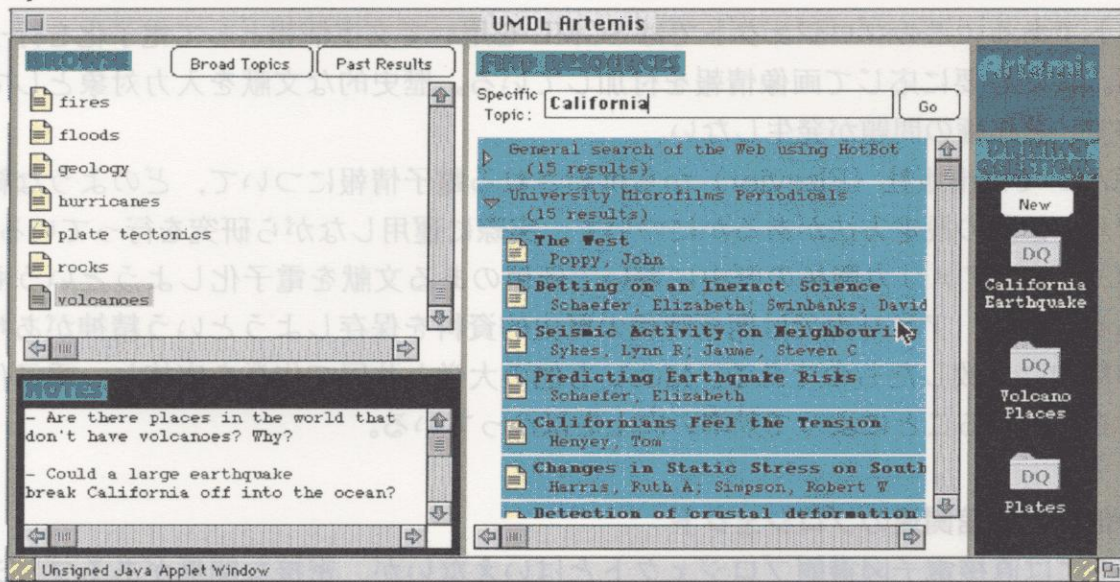


図 1.2 UMDL/DLI のユーザインタフェース

附属図書館プロジェクトとしての UMDL

The University of Michigan は総合大学であり、工科系学部のみならず医学部、芸術系学部などから構成されており、その歴史は 180 年を超えている。その附属図書館には豊富な資料が揃っている。その附属図書館も積極的にデジタルライブラリのプロジェクトを実施し、一部サービスを開始している。そのプロジェクトも The University of Michigan Digital Library プロジェクトという名前がつけられている。副館長の下にこのプロジェクトを実施する研究部門が中心となって、School of Information (情報学部) や ITD (Information Technology Division) と共同で研究を行っている。実際に行っているプロジェクトの詳細は次の URL から見ることができる。<http://www.umdl.umich.edu/>

附属図書館のデジタルライブラリプロジェクトでは DLPS (Digital Library Production System), JSTOR, Human Text Initiative (HTI), Making of America (MOA), PEAK (Pricing Electronic Access to Knowledge) など多数のサブプロジェクトが行われている。

JSTOR では、古い雑誌をスキャンすることによって電子化し、蓄積することを目的としている。古い雑誌を対象とすることで、著作権の問題を回避している。電子化の

対象としている雑誌は主に人文科学系の雑誌が多いようである。また、他の複数の大学・プロジェクトとも共同で行うことによって、負荷分散を実現している。さらに、契約を結ぶことによって、プロジェクトに参加していない研究機関でもデータが利用できる。(Media Union)

HTI では聖書やコーランなどの人類の歴史的にみて重要な文献を電子化するプロジェクトである。このプロジェクトでは SGML を用いて文字情報として電子化を行っている。また必要に応じて画像情報を付加している。歴史的な文献を入力対象としているので、著作権の問題が発生しない。

PEAK では出版社 (Elsevier) から供給される電子情報について、どのような経済モデル、値段の設定方法があるかについて、実際に運用しながら研究を行っている。

MOA は、アメリカ開拓の歴史において価値のある文献を電子化しようというものである。特に、アメリカでは開拓時代の歴史的資料を保存しようという精神があり、その精神に合致したものである。MOA も他の大学と共同で作業を実施し、電子化データを共有することによって効率の向上をはかっている。

その他の図書館関連のプロジェクト

以下には直接電子図書館プロジェクトとはいえないが、密接な関係があるプロジェクトについて紹介する。

CHICO (Cultural Heritage Initiative for Community Outreach) では、米国における黒人文化を Web を用いて紹介することを通して、学部学生に対して「どのように Web ページを作成するか」という教育を行うことが目的である。紹介しようとする対象について、どのようなことに着目して紹介するか、着目点をどのように紹介するかの方法論を、実際にコンテンツを作成することによって教育を行っている。技術的には特に新しいものを追求しているのではない。しかし、その内容は一見する価値がある。

プロジェクトなどをみて感じたこと

以上のようなプロジェクトが The University of Michigan では行われていたのであるが、これらのプロジェクトをみた上で感じたことを以下に述べる。

- 研究体制の日米間での違い

特に、図書館における研究体制の違いを実感した。日本の標準的な大学附属図書館には研究者がいない。これに対して、University of Michigan には研究者が多数在籍し、研究活動を行っていた。これらの研究者には博士号取得者が珍しくなく、関連する学部 (University of Michigan の場合は School of Information) において授業を担当していた。さらに、副館長のもとに研究のための責任者となる研究者がおり、この研究者のもとで附属図書館が行う電子図書館サービスのすべてを管理

している。もちろん、副館長とこの研究者は電子図書館研究も行っている。このような状況の違いを、在外研究期間の最後に日本の状況を Atkins 教授に話した時、教授は「Unbelievable !!」といていた。さらに、University of Michigan の体制は米国でもやや特殊な体制ではあるが、日本の状況は信じられないとも話していた。例え University of Michigan の体制が特殊ではあっても、実際に先端的な電子図書館研究を実施している米国の大学では類似の体制をとっており、日本の状況には深い危機感を覚えざるを得なかった。

● 研究指導者のリーダーシップの重要性

筆者の実質的受け入れ教官となってくれた Atkins 教授は、97 年当時学部長であった。彼は University of Michigan での電子図書館プロジェクトを率いており、その強力なリーダーシップにより研究が進められていた。筆者の目から見て、彼の指導力には卓越したものがあり、表現は難しいが、電子図書館プロジェクトに参加している研究者も彼の指導力を認めているように見受けられた。この点について、帰国直前に彼に尋ねたところ、次のような答えがかえってきた。「リーダーシップは才能だという人がいるが、自分はそうは思わない。努力により獲得できる。人の発言よく聞きくこと。そして、それを自分の中で消化し、自分の目標にあうようにうまくよりあわせていくことだ。」

また、School of Information の Cohen 教授にも Atkins 教授のリーダーシップについて尋ねたところ、次のような答えがかえってきた。「彼は優秀な（結婚の際の）仲人みたいだ。研究者が得意とする分野のテーマを、本人が気付かないうちに研究のテーマとして研究を始めさせている。」

University of Michigan は電子図書館の研究において優れた成果を着実にあげているが、優秀なリーダーの存在が大きいと感じた。今回の在外研究で得たもののうち、最も重要なものではないかと思う。

● 教育に対する取り組み

School of Information では先程紹介した CHICO などが行われている。CHICO は Frost 教授が中心になって行っている。彼女の場合は、「いかに Web の中で表現したいテーマを表現するか」を学生に具体的なテーマを与えてそれを実現することで、学生に教育を行っている。これは一見すると日本では専門学校などで行われているもののようにみえるが、取り組み方、奥の深さは比べ物にならず、研究としての用件を備えている。金物の技術としてではなく、ソフトウェアとしての技術の掘り下げ方には参考にすべき点が多い。さらに、今日のインターネットの普及、社会に及ぼす影響を考えると、このような教育を受けた学生が社会人として活躍することを考えると、我が国でも早急にこのような教育を行う学部、学科の整備が望まれるのではないかと印象を受けた。

2. ADL'99 に参加して

99年5月に米国 Baltimore にて開催された IEEE Forum on Research and Technology Advances in Digital Libraries 99 (IEEE ADL'99) に今井が出席し、科学技術論文における概念検索手法の発表を行った。今回の会議は Institute of Electrical and Electronics Engineers (IEEE) の Computer Society が Rutgers-CIMIC, GMD, National Library of Medicine, NASA-CESDIS と共催して開催されたもので、今回が6回目となる。この会議は5月19日から21日の3日間行われ、47件の投稿論文から16件の論文が採択され、発表された。セッションは一部2並列で、合計15のセッションが行われた。このうち、Keynote が3件（毎日1件の割合）、パネルが4件、ポスターが1件、全体講演が1件で、残りが一般セッションであった。Keynote として、

- Beyond the Endless Frontier: The Future of Research University in the Digital Age
- The Commercial View: Shipping the Digital Library V1.0
- Beyond Content-Based Retrieval: Modeling Domains, Users and Interaction

が行われた。このうち、一番目の講演者は University of Michigan の学長であった James Johnson Duderstadt であった。彼の講演の内容は電子図書館に関係したものではなく、今後の大学のあり方を彼の考えで述べたものである。彼の考えでは、現在米国にある数多くの大学すべてが研究と教育の双方を行う必要はなく、また非効率的である。限られた数の大学が研究も行い、他の大多数の大学は教育を行うべきであるという考えであった。会場からは University of Michigan のような研究面で現在高い評価を受けている大学を固定化してしまうのではとの発言が聞かれた。現在日本でも大学の改革が議論されている時期でもあり、興味深かった。

パネルセッションとして、

- Digital Earth
- The Publishing Industry and Generational Changes in Product Development
- Digital Sky
- Digital Libraries Initiative, Phase 2

が行われた。このうち、Digital Earth と Digital Sky については NASA が打ち上げた各種観測衛星からのデータなどを再構成して地球や大気圏の様子を仮想的に計算機の内部に構築しようというものであり、NASA 関係者がパネリストとなっていた。

NASA がこれまで電子図書館の研究に積極的に関与してきた理由の一つが、米国民の税金で行われている宇宙開発の過程で得られたデータを納税者に還元するための方策の一つであると考えられるので、ある意味で自然な研究の流れではないかと筆者は感じた。しかし、実際にプロジェクトとして実施されるとそのインパクトは強いものになるものとする。

一般セッションのセッション名は以下の通りであった。

- Ontology and Structure (3 件)
- Database Querying and Mining (3 件)
- Video (2 件)
- Mathematics and Music Applications (2 件)
- Text and Information Retrieval (3 件)
- Meta Data (3 件)

また、国別の発表論文件数は以下のようであった。

- アメリカ 8 件
- ドイツ 3 件
- フランス 2 件
- イギリス 1 件
- 日本 1 件
- シンガポール 1 件

やはり、研究が盛んなアメリカからの論文発表が全体の半数を占めていた。ついで多いのはドイツからの発表であった。ドイツの研究機関である GMD が主催団体に入っていたことも原因の一つであろう。シンガポールは小国ではあるが、情報産業の育成に力を注いでいることで知られている。このためか、以前からシンガポールでは電子図書館研究が盛んで、この種の国際会議において積極的に論文発表を行っている。これに対して、社会的関心が高く、電子図書館についての議論が活発に行われているはずの日本からの発表は筆者の 1 件のみという状況で、シンガポールのそれと比較して少ない感が否めない。これはこの会議での状況が特別なものではなく、電子図書館をテーマとした国際会議に共通する傾向である。日本での電子図書館に関する議論に参加する電子図書館を研究している研究者が少なく、技術の利用のみが注目を集め、期待されている現状を反映しているように思われた。

最終日以外は 2 並列でセッションが行われたため、筆者はすべてのセッションを聴講することはできなかった。各セッションとも 2 件から 3 件の発表が行われ、先にも

述べたように合計 16 件の論文の発表があった。筆者は Text and Information Retrieval のセッションで発表を行った。最近の電子図書館研究の一つの傾向に、メタデータの標準化とその応用がある。実際、メタデータ以外のセッションでもメタデータを直接的に取り上げた研究発表が行われ、メタデータの応用が始まりつつあることを示している。また、マルチメディアに関連した発表がいくつかのセッションにわけて行われており、ビデオ、音声、音楽などのマルチメディアを図書館でどう扱うかについての研究が広く行われていることが伺えた。また、オントロジーを用いた研究発表も行われており、字面の処理から内容に立ち入った処理の研究が始まっていることが感じられた。

これまで、ADL は 6 回開催されてきたが、2000 年は Association for Computing Machinery (ACM) が主催する The ACM Conference on Digital Libraries に合流することが報告された。実際、今回の会議の参加者は 100 名程度と少なく、特に最終日には参加者が少なくなってしまった。そのため、当初 2 つの並列セッションとして計画されていたものが、発表時間を半分に短縮することによって連続したセッションとして実施される程であった。開催場所が高級ホテルであり、参加費、宿泊費とも高額になったことも一因であろう。また、電子図書館という研究分野では情報検索 (Information Retrieval, IR) の研究分野の出身の研究者が多数を占めるようになってきているという現実を反映しているように筆者には思える。(IR の研究コミュニティの基盤は IEEE ではなく ACM である)

3. ECDL'98

1998年9月21-23日、ギリシャ、クレタ島ヘラクリオンにて、Second European Conference on Research and Advanced Technology for Digital Library(ECDL'98)が行われた。本会議はヨーロッパを中心とした電子図書館に関する発表、議論を行う会議であるが、アメリカ、アジアを含む世界各国から発表者、参加者が集まり活発な議論が交わされた。

セッションは最大4つが並行で行われた。Paper session, Spesial Session, Panel Session の他に DELOS Workshop が行われていた。今回の DELOS Workshop のテーマは電子商取引であり、電子図書館あるいは電子出版業界におけるビジネスに対する関心の高さが感じられた。

本学からは、デモンストレーション1件と、パネル1本の発表を行った。

- デモンストレーション

“NAIST Digital Library”, H.Sunahara, R.(Suzuki)Atarashi, T.Nisimura, M.Imai, K.Chihara

- パネル

“NAIST Digital Video Library: Tools for Restructuring Video Data for the Content Based Information Retrieval – A Representative Image of Shot Concept for the Internet”, Y.Kawasaki, R.Suzuki, H.Sunahara

デモンストレーションでは、会場から NAIST 電子図書館へアクセスするデモを行った。ちょうど台風が奈良を直撃、学内が停電した直後であり、その復旧の影響か多少反応が遅いこともあったが、大きな問題もなく終了した。実現方法などに関する関心が高く、多くの質問が寄せられた。

パネルはビデオ検索、シーン切替、検索など多方面からの質問があった。全体に評判は良く、実際にアクセスしてみたいという感想が多くあった。

ここでは、各セッション、パネル等のなかで参加したもの、重要なものについて報告する。Paper session については、技術的なセッションを中心に参加した。“Architectures for DLs”は各国の大学電子図書館で行われているプロジェクトの報告であった。アーキテクチャの概要の紹介や、各図書館の特徴となるサービスやアクセスの測定方法などの紹介があった。“Image DLs”は、イメージと動画を扱う技術に関する発表で、動画のシーン切替の自動切り出し、動画検索、またメタデータアプリケーションとしての応用などが発表された。“Multilinguality”ではスイス、日本、ギリシャなど非英語圏からの発表が続いた。質問のなかで、非アルファベット圏としての

大きな言語圏をもつ、中国(C)、日本(J)、韓国(K)の各言語の扱いに対する注意が特に喚起された。

他、参加できなかったが、Case Studies のセッションが3つあり、各電子図書館プロジェクトの報告が続いていた。

パネルセッションでは、”Metadata and content-based approaches to resource discovery”に参加した。Dublin Core を中心としたメタデータを専門するのスピーカと自然言語を専門とするスピーカとのパネルであった。どちらも重要な技術であるが、電子図書館の運用という視点では双方の技術をバランスよく取り入れることが重要と感じた。

4. Dublin Core Workshop

1998年11月2-4日、ワシントンDCのアメリカ議会図書館(Library of Congress)にて開催された、第6回 Dublin Core ワークショップに参加した。奈良先端大からは今回が初めての参加となる。Dublin Core はもともと WWW 上の文書オブジェクトのためのメタデータとして提案されたもので、基本的な目的は WWW 上にある情報資源の発見である。その後、より広い対象のメタデータとして利用できるように議論が進められてきた。現在、図書館の司書を中心としたグループに注目されている技術となっている。

Dublin Core(DC)の基本エレメントは15項目が提案されている。この15項目のエレメントセットは DC Simple(DC1.0)と呼ばれ、IETF(Internet Engineering Task Force) の推奨する標準化として、RFC2413 に登録された。15項目を以下に示す。

- Title: オブジェクトの名前
- Creator: 情報資源の内容に関して責任を持つ人または組織
- Subject: 情報資源に述べられたトピック
- Description: アブストラクトやイメージデータの説明など内容に関する記述
- Publisher: 情報資源を現在の形態にしたもの(出版社、大学など)
- Contributor: 著者ではないが文書の内容の作成に関わった人または組織(編集者や翻訳者な等)
- Date: 現在の形で利用できるようになった日付
- Type:
- Format:
- Identifier:
- Source:
- Language:
- Relation:
- Coverage:
- Rights:

図書館で扱う場合、これまでのメタデータとの互換性は必須である。DC では、複数のメタデータ規則を相互に変換できるしくみ Warwick Framework を提案している。

第6回 Dublin Core ワークショップでの議論の中心は2点にあった。1つは、DC1.0 の標準化をさらに進めること。もう1つは、Qualifier をつけることである。さらに、各プロジェクトでの実装を推進していくことが奨励された。