

データディクショナリ/リモートデータベース アクセス標準化とOSとの関連

溝口徹夫

三菱電機株式会社

ISOでのOSI標準化活動の中で、データベースに関するものの標準化が進められており、国内委員会による対応もなされている。本報告ではデータベース標準化活動の中でも、情報資源辞書システム(Information Resource Dictionary System:IRDS)とリモートデータベースアクセス(Remote Database Access)の2つを取り上げ、その最近の議論の内容を紹介しようとするものである。この2テーマは原案そのものが出来上がっておらず、概念レベルの議論、なぜ標準化が必要なのか、標準化の目的は何か、どのような表現が適切か、等論理の組み立ても含めて標準案を設定している段階である。我が国の貢献は現在の所それほど大きくないが、その内容を理解することによって建設的貢献が可能になる。また標準化の動向を紹介することで他の項目での標準化の対応の方法論や関連項目での方向付けにも多少の影響を期待するものである。

STANDARDIZATION ACTIVITIES ON DATA DICTIONARY AND REMOTE DATABASE ACCESS

Tetsuo MIZOGUCHI

MITSUBISHI ELECTRIC

In the area of OSI environment, recently standardization activities on database related items are initiated. Two of items addressed are Information Resource Dictionary System:IRDS and Remote Database Access:RDA. Within OSI framework, database is related closely to other components when we start talking about distributed information resource management and distributed database access.

Recent gatherings in Washington and London produced documents to proceed to Working Draft hoping to finally be International Standards.

The present paper gives the progress reports how drafts are taking shape and what was the argument on such specific proposals.

0. 目次

1. データベース関連の標準化概要

IRDS/RDASPの位置付け

2. 情報資源辞書システム(IRDS)の標準化 範囲、目的、原案の内容、問題点

3. リモートデータベースアクセス(RDASP)の標準化 範囲、目的、原案の内容、問題点

4. OSとの関連

0. 1 はじめに

本報告はデータベース関連の国際標準化の動向のなかで特にオペレーティングシステムと関連の深いものについて述べるものである。

標準化の議論のなかにはいろいろの判断基準があり、特に政策的なものと技術的なものが混在し得る。本報告は技術的問題を扱うものである。標準化活動の内容は今後合意の基に決定していくものであるから、十分な完成度に至っていない状態のものが多く、従って国際のレベルでどこまでなにか正式に取り決められたか、どのような可能性が今後有りえるのか、国内での対応／考え方はどうか、個人的見解はどのようなものか(ここで言う個人とは国際委員会に参加している専門家を指す)は流動的であり、個人的見解と国内委員会の見解等を明解に区別出来ないこともある。本報告では出来るだけ区別を明解にすることにする。

他の活動がそうであるように、標準化活動には次のことをはっきりさせておかねばならない。

- (1) 標準化の範囲をどのように設定するか
- (2) 標準化の目的は何か、必要性は何か
- (3) どのような表現方法が妥当か
- (4) 標準化の内容

以上の点を可能な限り本報告では述べてみたい。これらのことは一般に標準の文書には充分述べられていないことが多い。

とは言え、十分な理解の上での報告であるか否かについては疑問が多い。また個人的価値判断も当然入っている。それらの事情を理解された上で本報告を御検討頂くことを御願いたい。

1. データベース関連の標準化概要

ISO/TC97の中に開放型システム(OS: オープンシステム; オペレーティングシステムではない)の標準化のための委員会SC21があり、またその中がいくつかのワーキンググループWGに分かれ、その一つとしてWG3がデータベースに関する標準を扱っている。国内でも同様の組織化が成されている。昭和61年4月現在扱っているものは以下のものである。

(1) データ管理参照モデル

(2) データベース言語 SQL

(3) データベース言語 ND L

(4) 情報資源辞書システム IRDS

(5) リモートデータベースアクセス RDASP

特に後者の二つは最近追加されたものである。SC21の中においてデータベース標準化はデータベース言語のみを扱っている間は孤立した活動であったが、IRDSやRDASPの追加で他のWGとの接点が出てきた。このような接点での他のWGとの協同作業と共に、WG3内の一貫性の保持も必要になりつつある。その例をいくつか述べると、

- a) IRDSでは辞書情報をアクセス出来るようにSQL, ND Lとのインタフェースを準備するのは当然であるが、IRDSの機能を実現するための仕様記述言語としてSQLを使う試みがある。
- b) RDASPでの対象とするデータベースとしてSQL, ND Lを考えに入れなくてはならない。他のWGでコミットや障害回復の機能を汎用的に提供されることが決定されつつある(CCR)が、データベース言語との整合性を保証しなくてはならない。

ここで少しOSIの階層のことについて触れる。既に広く知られているように、ISOでのOSIの階層には7層あり、最上位が応用層である。通信という側面から見た応用層のプロトコルとサービスとはどのようなものであるかを理解するのは困難を伴うものではあるが、応用層には既に述べたCCR等の共通サービスを提供するCASE等のものも含まれる。データベースはOSIの階層では第7層に位置付けられている。データベース単体としては、直接には通信と係わらない部分も多く、従って応用層にあってもなんらOSIとは関連を持たないで済むものもあるが、本報告で触れるのは周辺との関連が出てくると思われる2つの項目についての標準化の話題である。

以下では特にIRDSとRDASPにのみ話しを限定することとしたい。

国際標準には、年1回のWGでの総会による各国代表の意見の集約と標準決定の答申案の上部への提出の他に、専門家の会合による具体的な標準の内容の検討とが随時行なわれる。

このような作業が開始されるのは標準として取り上げる項目としてふさわしいか否かの決定後、新たにグループをWGの中に作ることで始まる。

最近のIRDS, RDASPの活動は以下の通り。
IRDS 昭和60年11月 ワシントン
実質的な第1回専門家会合兼WG3総会

昭和61年3月 ロンドン

専門家会合

RDASP 昭和60年11月 ワシントン

新たなグループの編成を総会で認可

昭和61年3月 ロンドン

第1回専門家会合

本報告はIRDSについてはワシントン、ロンドン
RDASPについてはロンドン会合を基にしている。

2. 情報資源辞書システム (IRDS) の標準化 Information Resource Dictionary System

2.1 標準化提案の歴史的背景

データについてのデータ (メタデータ等と呼ばれる) をデータベースではデータ辞書と呼び、データベース管理システム (DBMS) と組み合わせて使用されたり、データ辞書単独で使用されてきた。

データ辞書はどのようなものであればよいか等についてはいくつかの調査研究がなされ、特に英国計算機学会 (BCS) での調査報告が ACM・SIGBDP に掲載されている。

一方米国内で、特に米国政府内での、データ辞書の標準化によるデータ処理の効率化が叫ばれ、政府関係のデータ辞書標準作業が行なわれ、FIPSとして承認されるに至った。米国NBSを中心としたこの活動が米国内でANSI標準案、更にはISO国際標準案として提案されるにいたった。

この標準案が正式に各国に送付されたのは、昭和60年X月である。この文書は千ページにも及ぶもので、概説資料付である。

この文書についての各国からの意見を基に、昭和60年11月ワシントンでのIRDS専門家会合で具体的な検討に入った。

IRDSの討議の内容を以下のまとめ方で述べる。

- ・標準原案の要旨
- ・標準原案への各国の反応と代替提案
- ・代替案の内容と標準原案との比較
- ・今後の発展

2.2 標準原案の要旨

ANSI原案 (dpANS : draft proposalANS) の基本的考え方は、政府データ処理部門に使用法の違う種々のデータ辞書システムが存在することは効率上好ましくなく、この利用者インタフェースを標準化することで多くの費用の節減になるというものである。従って標準化の作業もまず既存のデータ辞書の機能を調査分析し、辞書に格納されるデータの種類を抽

象度で分けて、更に利用者インタフェースを完全に決めたものである。

dpANS の特長を挙げてみると、

- (1) データ辞書を利用者インタフェースという観点から捉えている。従って標準の記述は全部コマンドの列挙になっている。
- 原案著者の見解はデータ辞書の機能を表わすにはコマンドで固定的に (concrete) 決めなければ標準の意味は無いし、製造関係者はこのような仕様に満足しているというものである。
- (2) データ辞書というソフトウェアツールの標準であってデータ辞書の方式 (アーキテクチャ) について言及するものではない。

このことはこの原案は孤立したツールとしてしか使えないのではないかと印象を与える。

方式についての考え方が充分読み取れない、また膨大な資料をしかもコマンドでの表現をたどって理解することは不評を買った。

- (3) データ辞書の対象はDBMSである必要はなく、またデータ辞書の機能はDBMSのデータ記述の機能とは重複する部分もあるが、かなり独自の機能が必要である。

このこと自身は正しいことであるが、同じ標準化活動を行なっているSQL, NDLの配慮がない。

- (4) コマンドインタフェースの標準が主体であるが、データ辞書を設計するときのモデルがなければ枠組みがはっきりしないということでER (Entity-Relationship) モデルを採用している。

ERモデル自身の簡便性は認められるが、本モデルには異論を唱える向きがある。

(モデル論を持ち出すとどのようなモデルでも不十分であることは仕方ないかもしれない)。

2.3 dpANSの審議：昭和60年11月 ワシントン：標準原案への各国の反応 と代替提案

dpANSに対して議論に参加したのは、米、英、仏、加、日であった。米以外のdpANSに対する評価はかなり低く、加と日の意見は偶然にも似通ったもので、ともに消極的なものであった。特に英の意見は目的がはっきりしない、アーキテクチャがきちんとしていない等厳しいものであった。

以下にそのコメントを挙げてみる。

- ・IRDS標準化の目的を明記すべきではないか。
- ・ISO 標準のSQL, NDL, IRDSは統合すべきか。
- ・IRDS利用者は標準のDML でIRDS定義情報をアクセス出来るようにするか。

- ・エンドユーザインタフェースはIRDS標準に含むべきか。
- ・IRDSの機能性はconcreteな形式で表現するか、またはabstractな形式にするか。
- ・制約条件は出来るだけ宣言的に表現するか。
- ・ISO 標準のERモデルが1個だけとするか。
- ・IRDS標準の一部としてEAR モデルを考えるか。
- ・IRDSはactiveなデータディクショナリシステムか。
- ・IRDSはRDASP を含むのか。
- ・唯一のIRDSがSQL,NDL が同時に共存する環境をサポートできるとするのか。
- ・システム標準スキーマは最初のIRDS標準に入るのか。
- ・最小システム標準スキーマはSQL,NDL の概念を含むのか。
- ・IRDS情報とIRDSスキーマ情報の間に関連を定義できるか。
- ・スキーマは自己定義可能か。

以上のコメントとその背景の議論が行なわれた後、次の3項目について整理を行なうことで合意した。

- (1) IRDS標準の目的
- (2) IRDSアーキテクチャの検討
- (3) 表現方法

つまりここで代替案がd p A N Sとは独立の枠組みで、英国の主導のもとになされることとなった。

IRDS標準の目的：

各国の意見をベースとして目的を明確にした。

IRDSアーキテクチャの検討：

標準原案に不足していたアーキテクチャの整理とデータ管理参照モデルとの整合によるIRDSの枠組の提案を行なった。

表現方法：

コマンド形式ではなく、出来るだけ宣言的に条件の表現を行なう。表現言語としてSQL を使ってみる。

以下にその詳細の一部を示す。

3.5	Modules of the IRDS standard shall define the schema extensions required, using the extensibility facility, and functionality to support the ISO standard:		・ NDL, SQL, ISO standard programming languages ・ Open System Management and Directory Services
	STATEMENT OF IRDS OBJECTIVES FOR INCLUSION IN ISO STANDARD		
1.0	<u>SCOPE</u>	4.0	<u>SHARING IRDS DATA</u>
1.1	An IRDS shall be able to provide a shareable repository for all definitions and descriptions needed, produced, and/or used during the analysis, design, construction, and use of a system.	4.1	An IRDS shall provide users of ISO standard database management and programming language facilities with a common means of specifying data definitions.
1.2	A single IRDS shall be able to provide support as described below for all ISO standard database languages and programming languages as described in the following objectives.	4.2	The common data definitions maintained by an IRDS shall be accessible to programming language compilers, database management systems, and any other application through one or more standard interfaces.
2.0	<u>REFERENCE MODEL</u>	4.3	An IRDS standard shall not preclude an integrated implementation with a DBMS based on ISO standard database language NDL or SQL.
	The IRDS shall be consistent with the Reference Model developed by ISO/TC97/SC21/WG3.		
3.0	<u>MODULARITY AND EXTENSIBILITY</u>	5.0	<u>ACCESS TO IRDS DATA</u>
3.1	The design of an ISO standard IRDS shall be such that an initial standard can be issued and further modules developed, possibly by ISO working groups other than ISO/TC97/SC21/WG3, at a later time, to support such things as the following:	5.1	An IRDS shall be able to provide facilities for programs written, using ISO standard programming languages, to use the SQL Data Manipulation Language to access information stored in the IRDS at execution time.
	a) documentation b) system life cycle and project management c) data element standardisation and management d) organisation planning e) support of database administration, document administration, information resource management and data administration f) support for distributed processing and database environment g) a source and object library management system h) quality assurance i) configuration management	5.2	An IRDS shall be able to provide facilities for programs written, using ISO standard programming languages, to use the NDL Data Manipulation Language to access information stored in the IRDS at execution time.
		5.3	An IRDS shall allow access to its data from another Open System using standard OSI protocols made using extensibility from one IRDS to another IRDS.
3.2	The IRDS standard shall be modularised so as to support a wide range of operating environments and to support cost effective procurement.	6.0	<u>TRANSFERABILITY TO IRDS SCHEMA CONTENTS</u>
3.3	The IRDS standard shall define a minimum or basic schema required to support:	6.1	It shall be possible to transfer all or part of the IRDS Schema, including any extensions made using extensibility from one IRDS to another IRDS.
	・ extensibility ・ integrity ・ access control ・ auditing	7.0	<u>TRANSFERABILITY TO IRDS DICTIONARY CONTENTS</u>
		7.1	It shall be possible to transfer all or part of an IRDS's dictionary data from one IRDS to another IRDS.
		8.0	<u>PORTABILITY OF SKILLS</u>
3.4	The IRDS shall provide the capability to specify the types of information to be held by the IRDS and the constraints to be imposed thereon.	8.1	Users of multiple standard-conforming IRDS's should not need significant retraining. This requires specification of standardised content in a Panel Interface and the option of an concrete command language.

2. 4 代替案の内容と標準原案との比較

既に2. 2の標準原案の要旨の説明で部分的に述べているものをここでまとめてみる。

IRDSの目的

標準原案：統一的ソフトウェアツールの提供

代替案：枠組みの設定と目的の明示化によるソフトウェアツールの標準

標準の表現形式

標準原案：コマンドの規定による詳細な標準設定

代替案：データ記述の充実による表現

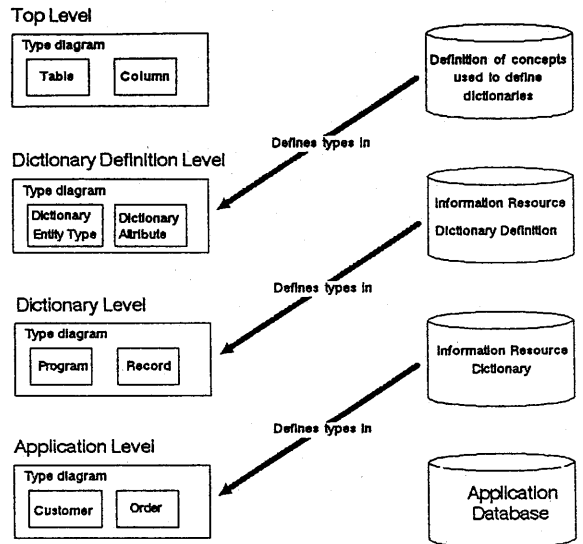
この2点が特に際立った相違点である。

標準原案がISOに出された後ANSIの中でIRDSの改訂作業が行なわれている。とくに大きな変更点は核になる部分を少なくし、モジュールとして外付けにするものが多くなった点であるが、討議の中での標準原案への質問に対して、“それはANSI改訂版では訂正した”との発言が目立ち移動目標を追っているとの文句も出ている。

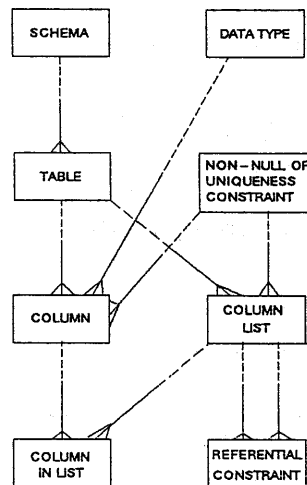
一方標準原案を作った側から、標準を構築する側から見れば、標準は詳細かつ具体的であるべきで、そのためにはコマンドによる機能性の側面から望むべきであるとの意見が出されている。この意見に対する1つのコメントは確かにその通りであろうが、千ページにも及ぶ文書を複数の、多分10人以上の技術者が実現をしようとするとき、全体像の見えないものだと混乱を招くのではないかというものであった。

もう1つ大切なことは実現のしやすさよりも、利用者がIRDSの概念を性格に把握できなければならないと言うことである。IRDSの利用で最も大変なのは、概念や方法論であって、ソフトウェアの使い方は二次的ではないかと考えられる。

標準原案を作った側からの注文としては、代替案ではIRDSの機能性を表現出来ていないということであった。昭和61年3月のロンドン会議は代替案の最終審議ではなく、提案の方向性についての意見を集めるためのものと性格付けられていることから、不十分さが存在することは認めざるを得ない。



Levels and the IRDS



Top Level data

2. 5 今後の発展

既に述べてきたように、IRDSには現在2案あり、その選択を昭和61年9月のロンドンでのWG3会議で行なう予定である。

米国は標準原案をそのまま認め、代替案を技術報告として添付するよう求めており、他の国は日本を含めその要求は認め難いとしている。

代替案の完成を待って、あるいはその内容を洗練化することが当面の我々の採るべき立場と考えられる。

日本としては更に次のような立場を明らかにしており、その立場は今後も主張すべきと考えられる。

IRDSは次の3点の見方ができる。

- (1) IRDSを文字通り“情報資源管理”と広くとらえ、情報のモデリング手法、情報システムの構成法、ソフトウェア管理パッケージ等を含むものとする。
- (2) IRDSは少し狭く“ディクショナリ”と考え、情報資源も自動化されたものの表現を主体に考える。
- (3) OSIの環境での分散等に関する資源管理について考える。

標準原案は(2)にのみ関するもので、それに対する各国の意見は(1)も考慮に入れるべきというものである。(1)をどこまで標準とすべきか(特にOSIとして)には大いに議論のある所である。

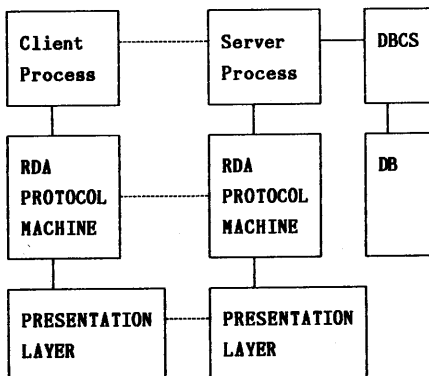
3. リモートデータベースアクセス(RDASP)の標準化 Remote Database Access and Services Protocol

RDASP の話を次のようにまとめてみる。

- ・標準原案はどのような主旨のものか
- ・標準原案への反応
 - 各国の反応、提案者の改善提案、環境の変化
- ・将来予想される方向
 - この3点を以下に述べる。

3.1 標準原案(SC21/WG3/N41)の主旨

ICL社G.Brown を中心にECMAの委員会で3年前から検討されていたテーマであった。



図に示すClientProcess からServerProcess へのDBアクセス要求がOSI 第7層に位置するRDAProtocolMachineを経由して実行される。第6層以下の存在の上にRDAServiceとRDAProtocol を規定しようとするものである。

提案の特徴をいくつか挙げてみると、

- ・ワークステーション-ホストの関係を意識した提案であり、ホスト/サーバは1個の考え方であるが、複数ホスト/サーバにも拡張をほのめかしている。
- ・DBはSQLを基本に検討されており、SQLを一部拡張している。例テンポラリテーブル。
- ・コミット/リカバリ機能を持っている。
- ・データ転送はDB言語とのインタフェースとの関係でタプル(レコード)単位で行なわれる他、バルク転送機能を追加している。
- ・コマンドの通信回数を減らすマクロ機能や同期の役目に使うコマンド実行のグルーピング機能を持つ。

標準原案は明らかにいくつかの矛盾点と課題点を含んでいる。これらは原案への各国からの反応という形でまとめてみる。

3.2 標準原案への反応

(1) 各国の反応

- 各国と言っても、意見を出したのは西独と日本である。以下にその要旨を示す。
- ・RDAで使っているSQLは標準SQLと互換でない。
- ・プロトコルはもっと柔軟なものであってよい筈。
- ・RDAではDB言語の機能性に依存すべき、またはみだした機能は付けるべきでない。例マクロや複数レコードアクセス。
- ・小規模システムにも使えるよう2レベルの設定をしてはどうか。
- ・RDAと分散DBのための核との差がはっきりしない。RDAの標準の範囲をはっきりさせたい。
- ・標準でのSQL, NDLでのモジュールとRDAのマクロ等の関係が付いていない。
- ・スキーマ情報をどのようにアクセスするか。
- ・RDAでレコード単位の転送は適切ではない。
- ・SQLでは1フェーズコミットでRDAでは、2フェーズである。
- ・原案ではワークステーションとホストの1対1の構成しか考えていない。

(2) 提案者の改善提案

提案者は問題点指摘を予期して、RDASP の利用側面を分かりやすくするためのシナリオ作りを提案し、例えばリモートとローカルDBの関係付け等について明らかにしようとするものである。

(3) 環境の変化

原案提案者によれば、RDASP 検討開始時期と比べて現在では、

- ・応用層にCCR 等のCASEの標準化が進んだこと、
 - ・OSI への提案という形をとっていないがCCITT のMHS に源を発するROS(Remote Operation Service)の採用の動きがある(ECMA)、
- ことである。

3. 3 将来予想される方向

- 昭和61年3月のロンドン会議での合意事項は
- ・ROS を基本とした形にRDASP 原案を書き換える。
 - ・RDA ではデータベーススキーマへのアクセス機能が必要である。
 - ・テンポラリテーブルの利用は有効である。
 - ・ASN.1 を用いてデータベースデータとプロトコルの記述を行なう。

等であり、編集者への宿題は、

- ・利用のシナリオの文書化、
 - ・バルクデータ転送への要求を整理する、
 - ・バルクデータの効率的エンコード法を提案する、
 - ・基本(generic)DMLタイプをはっきりさせる、
 - ・マクロの扱いはっきりさせる、
- ことである。

(RDASPについての記述はこれで終るつもりであったが、編集者より新原案が届いたので概説を追加する)。

(1) サービス機能

RDA(Remote Database Access) でのサービスにはいくつかのオペレーションがあり、そのオペレーションは次の4種に分類される。

- ・アソシエーション管理
クライアントとサーバ間の関係管理
CASE核へ写像
- ・トランザクション管理
CCR へ写像

・データ操作

単一DML 文の実行制御

ROS へ写像

・繰り返しデータ操作

同一DML 文の指定回数繰り返し実行制御

ROS へ写像

具体的には次のオペレーションが提供される。

・アソシエーション管理

R-Connect, R-Release 他にR-SetStatus

・トランザクション管理

R-BeginTransaction, R-PrepareToCommit,

R-Commit, R-Rollback, R-Restart

・データ操作

R-ExecuteDML

・繰り返しデータ操作

R-PrepareRepeat, R-ExecuteRepeat

以上の構文はASN.1 を用いて記述されている。

以上のオペレーション以外にデータ表現が記述として含まれている。

(2) RDA の動作状態

一つのRDA コネクションでは単一のクライアントが単一のデータベース上で動作する。複数のデータベースを動作させるには、別途RDA コネクションを張る必要がある。

一つのRDA コネクションでの動作は状態によってモデル化出来る。

状態名	可能な動作
Idle	R-Connect * (C) R-SetStatus * (C)
Connected	R-BeginTransaction * (T) R-Release * (I)
TransactionIdle	R-ExecuteDML R-Commit * (C) R-Rollback * (C) R-PrepareToCommit * (S) R-PrepareRepeat * (R)
Secure	R-Commit * (C) R-Rollback * (C)
Repeat	R-ExecuteRepeat R-EndRepeat * (T)

ここで*印は状態の変化を示し、括弧内の文字は状態の頭文字を示す。印の無いものはその状態で発行できる。

(3) RDA の占める位置

既に述べたようにRDA はCASE核、CCR、ROS の上に実現するものとして位置付けられる。

RDA 機能として単一文実行とコミット制御は必須であり、マクロ機能と繰り返し動作はオプションであるとしている。

4. OSとの関連

ここで言うOSとは何かは深く追及することは止めて、IRDS、RDASPの周辺との関連を考えてみる。
IRDS

前述の如くOSI の中でもDirectory Service というものがあり、分散環境での資源管理には大いに関連を持つものと考えられる。但し現在のIRDSは論理情報に主たる問題設定をしており、分散環境固有の問題を扱っている訳ではない。自ずからIRDSとDirectory Service の役割分担がなされるのであろうが、Directory Service はCCITT の世界からの情報のモデル化を行ない、IRDSは情報処理の世界からのモデル化を行なっており、考え方に違いがあり、それを統一すべきか否か、一考を要する。IRDSはDirectory Service よりも広く(と筆者は認識しているが)情報資源の管理を行なうことを目指しており、各種のインタフェースを持ち得ることから、今後更に複雑な機能性へと足を踏み入れると、単にDirectory Service との関連だけでない処理や運用一般の問題へと課題の対象が広がる恐れがある。

RDASP

RDASP はOSI の枠組みの中に入りこんだものであることから、少なくともOSI 内での各要素との関連について具体的検討がなされ得るものであろう。各要素を接続して動作させるものとしては、事例は多くない現状ではどのような問題があるか、RDASP 自身の完成度も充分でない現状では今後の方向を予想することは難かしいが、応用層の構成の具体的提案としての意味は大きく、この延長線上での検討が今後進み得る。例えば分散DBとの関係ではどのような違いが分散DBとRDASP の間にあるのか、その差の為に必要な要素や機能は何か、等具体的議論へと発展し得ると考えられる。