

特集にあたって

吉川 正俊 (名古屋大学)

XML (Extensible Markup Language) は、テキストによって半構造化データを表現するためのメタ言語である。Web 関連の業界団体である W3C (World Wide Web Consortium) によって 1998 年に標準が勧告され、その後も、プログラムからデータを操作するための API (Application Program Interface) や、XML 文書の一部を宣言的に指定するための XPath 言語、XML 文書の文法やデータ型を詳細に記述するための XML スキーマなどの標準が相次いで勧告された。XML は、特定のハードウェア、オペレーティング・システム、通信プロトコルに依存しないことや、それぞれの応用に適した言語の文法を定義することができるメタ言語であることなどの特徴を持つため、データおよび文書を表現するための汎用の基盤技術として広範な分野で利用されている。XML の応用分野をすべて列挙することは到底不可能であるが、ごく一部の例として、ゲノムデータなどの科学技術、ニュース配信やデジタル放送などの情報配信、電子商取引や電子政府を実現するための各種電子文書などを挙げることができる。

大量の XML 文書が、データ交換、フィルタリング、蓄積に利用されると、当然、データベースとの関連が重要になる。XML データベースは、近年、理論、実際の両面で極めて活発な研究が行われている分野であり、また、多くのベンダによって商用システムも開発されている。本特集では、このような XML データベースの現状を俯瞰することを狙った。最初の 4 編は、国際的に活躍している気鋭の研究者による個別テーマの解説であり、続く 2 編は、ベンダの方による最先端の商用システムの解説である。まず、天笠、吉川による「XML データベース技術概説」では、XML の簡単な紹介に続き、大量の XML データを格納し、高速検索、更新機能を提供する XML データベースを実現するために提案されている代表的な手法とその基礎技術を解説している。次の田島による「XML 用木パターン言語 XPath 解説」で解説される XPath は、

XML 関連の標準体系の中で重要な位置を占めている。例えば、XML のための問合せ言語 XQuery は XPath を基礎にしており、XML データの変換言語 XSLT でも XML 部分文書の指定には XPath を用いる。また、岩井原による解説にもあるように、XML 文書のアクセス制御においても XPath が対象部分文書の指定に用いられる。田島の解説では、XPath を評価するための時間計算量や並列化可能性などに関する最新の研究成果を紹介している。さて、天笠、吉川が、大量の XML データをデータベースで管理する手法を論じているのに対し、森嶋による「XML 出版の技術」では、関係データベース内などに既に大量に存在する非 XML データを所望の書式の XML データに高速変換する XML 出版について解説を行っている。さらに、岩井原による「XML 文書のアクセス制御」は、XML を重要な社会システムに展開する上で必須の技術を解説している。XML のアクセス制御やセキュリティに関する標準化は、電子ビジネス関連の標準化を推進しているコンソーシアムである OASIS で進められている。岩井原の解説では標準化と研究の最新動向を紹介している。最後 2 編の菅原による「IBM 社のリレーショナル・データベース DB2 における XML データの実装方法」と鈴木、秋本による「Oracle XML DB に見る統一スキーマの実現と課題」は、ベンダによる商用システムの紹介である。データベース管理システムは巨大なソフトウェアシステムであり、既存の機能と整合性を取りながら新たな機能を付加することは容易なことではない。ベンダによって着目点と利用している技術が異なることが興味深い。

XML データベースは、広範な応用分野であり、したがって大きな市場を持つ。過去長期にわたって蓄積されてきたデータベース技術に立脚した学問的な研究成果が商用システムに影響を与えることが多く、シーズとニーズが相互に刺激しながら発展している技術分野である。本特集がこの分野の現時点でのスナップショットを伝えることができれば幸いである。