

XML を用いた技術史の知識基盤表現 Knowledge Representation of Technology History Using XML

松井 正志・中平 勝子・三上 喜貴

MATSUI Masashi・NAKAHIRA T. Katsuko・MIKAMI Yoshiki

長岡技術科学大学

Nagaoka University of Technology

知識表現, XML, 年表, 視覚化

Knowledge Representation, XML, Timeline, Visualization

1. 目的

日本の技術革新に関する知識基盤を構築するためには、技術史に関する知識の記述について、表現力が高く、再利用性に富んだ記述形式を開発することが必要である。また、視覚化と統合化により、蓄積された知識を直感的にわかりやすい形に表現するツールの開発も必要である。本稿では、技術史を Web 上で知識基盤化するにあたり、どのような技術史記述の枠組みが必要かについての検討を行った。

2. これまでの研究

これまでに国内外を問わず、歴史的知識を記述するための多数の試みが行われてきた。国内では、電気学会の電気技術史技術委員会が1996年にハイパーメディア作成のためのワーキンググループを組織し、電力系統技術歴史データベースの作成を行った¹⁾。当時はまだXMLが開発されていなかったという背景もあり、独自の歴史記述言語 HSML (Historical Space Modeling Language)²⁾を開発し、データベースを構築した。また、データの検索と表示のGUI (Graphical User Interface)として「曼荼羅」というインターフェイスが開発された。このインターフェイスによってイベントの遡及、波及関係をたどることで経時的推移をたどることができる。海外ではHEMLプロジェクト³⁾が歴史記述を目的とした独自のXMLスキーマと、そのデータをSVGでグラフィカルに表現するWebアプリケーションを開発している。

一方、AJAX (Asynchronous JavaScript + XML)に代表される近年のWeb技術の発達に伴い、歴史に関する情報を統合的に表示できる様々なWebアプリケーションが開発されている。例えば、歴史的な事実の集合を二段階の時間軸目盛を持つ年表形式で表示するTimeline⁴⁾がMITなどをメンバーとするSIMILEプロジェクトによって開発された。

表1 . HSML, HEML, Timeline のタグの比較

	HSML	HEML	Timeline
分類	Classの継承により階層別分類が可能	記述可	なし
時間	年(月,日)で指定	年(月,日)で指定	年(月,日,時刻)
場所	明示的にはなし(但し属性の一部として記述可)	記述可	なし
関与者	開発者を記述可	記述可	なし
参照情報	記述可	記述可	記述可
他の属性	多数の属性を記述可	なし	表示アイコン

3. イベント記述枠組みの検討

技術史記述のための枠組み検討の第一段階として、まずイベント(歴史的イベント)記述用タグを検討した。検討素材としてHSML, HEML, Timelineを取り上げ、そのタグ構造について比較した。HSMLについては、適用分野に応じてタグ構造を拡張して使用することができるが、ここでは開発者が示しているソフトウェア向けのタグ構造を用いた。表1に示すように最も表現力が高いのはHSMLであるが、一般的な歴史的イベントの記述に対してはタグの種類が細部にわたり過ぎる部分もあり、年表記載事項程度の記述にはHEMLで十分と思われる。次にタグの表現力や使い易さを実例で確認するため、歴史記述にHEML, 表示形式にTimelineを使い年表の試作を行った。試作のための歴史データは、電気学会電気技術国産化の歴史調査専門委員会などの実例を基にした^{5), 6)}。表示形式にHEMLの提供しているWebアプリケーションを利用しなかったのは、HEMLのXMLスキーマの相互利用性を検証するためである。試作した年表が図1である。



図1. HEML/Timeline を用いた年表表示の例 (下段の少し色の濃い帯は10年を単位とする目盛)

4. イベント間の関連記述枠組みの検討

第二段階として、イベント間の関連を記述する方式について検討した。技術史に関する記述は、イベントの属性に関する記述のみならず、イベント間の因果関係をはじめとする関係記述が必要と思われるが、実際、技術史に関する文献の中で、技術とそれを取り巻く環境や社会的背景との関連に関する記述がどの程度の比率を占めるかを確認するために調査を行った。その際、技術史文献の中でも、著者や扱っている分野によって傾向が異なることが考えられるため、技術的記述が中心となると思われる文章として電気学会電気技術国産化の歴史調査専門委員会の報告書⁵⁾を用いた。また、社会的背景などに大きなウェイトをおいた文章として、トレヴァー・I・ウィリアムズの「二〇世紀技術文化史」から「家庭における技術」⁶⁾を用いた。

調査結果は、表2のとおりである。電気学会報告書では約60%の文が技術自体に関する記述であるのに対して、「二〇世紀技術文化史」では技術自体に関する記述は約40%である。一方、両文献において、社会経済環境に関する記述は15%~25%を占めており、イベント記述の枠組みは技術的な内容以外のイベントも考慮する必要があることを示している。また、これらのイベント間の相互関係に関する記述は10%~30%を占めており、文献の性格によらず、イ

イベント間の関連記述が相当程度必要であることを示している。

そこで、イベント間の関連記述の枠組みをどのように設定するかが課題となる。この点、HSMLはイベント間の関連についてもイベントの属性として記述する方式をとっており、先行イベントや後続イベントなどを指示するタグがある。その他、`used_by`、`developed_by`などのタグが用意されており、更に必要な関連記述タグを独自に定義することができるなどの工夫により、極めて高い表現力を実現している。しかし、これらのタグを使用するためには人物や組織といったイベント以外のクラスも導入する必要があり、記述体系が複雑とならざるを得ない。

5. 今後の課題

技術史文献の記述内容の分析結果が示すとおり、イベント間の関連記述は必須である。本稿で取り上げたHEMLは年表記述には十分であるが、イベント間の関係を記述することができず、他方HSMLは関係記述が可能であるが、タグの種類が豊富すぎてデータ作成の負担が過大となる懸念があるなど一長一短がある。今後、技術史の関係記述の類型化についての調査と考察を進め、これにふさわしいXMLスキーマを開発することが課題である。

表2. 技術史文献における記述内容の分布

(1) 電気学会報告書		(2) 二〇世紀文化史	
項目名	文章数	項目名	文章数
技術	50	技術	31
技術から技術	1	技術から技術	7
技術から環境	3	技術から環境	7
環境	11	環境	20
環境から環境	0	環境から環境	2
環境から技術	4	環境から技術	9
その他	13	その他	3
合計	82	合計	79

参考文献

- 1) 電気技術史技術委員会, 電気学会技術報告, 第991号, pp. 9 - 11
- 2) Matsumoto, Y., A. Yamada, An association-based management of reusable software components, Annals of Software Engineering 5 (1998)
- 3) Robertson, B.G., Visualizing An Historical Semantic Web with HEML, WWW2006, May 2006, Edinburgh.
- 4) SIMILE Timeline, <http://simile.mit.edu/timeline/>
- 5) 電気学会電気技術国産化の歴史調査専門委員会 (1999), 技術創造, 朝倉書店
- 6) 高橋 茂 (1996), コンピュータ クロニクル
- 7) トレヴァー・I・ウィリアムズ (1987), 二〇世紀技術文化史 下