

# 文字もじMOJIの世界

## 13. COBOL資産活用における文字情報基盤への対応

河合 孝志\*

Society5.0（官民データ連携）実現に向けた「デジタル・ガバメント」の取り組みの一環として政府では、「IMI（共通語彙基盤）」を共通データ標準としたデータ連携基盤の構築を加速している。その中で「文字情報基盤」の整備にかかわる取り組みとして、約58,000の漢字の文字セットを国際規格として成立させた。そこで浮上しているのが、既存の「COBOL（Common Business Oriented Language：共通事務処理用言語）資産」をそれにどう対応させていくかという課題である。

長きにわたり企業システムを支えてきたCOBOL資産は、今日も保守・改修が加えられながら広範な分野のビジネス活動を担うシ

ステムとして稼働している。そうした意味でCOBOLは、決して“過去の遺産”ではなく、まさに“現役”として今も活躍を続けている。

外的環境変化（政治面、経済面、社会面、技術面）へのCOBOL資産の扱いにおいて、今日とられているアプローチは大きく3つある。1つめは最新の技術との連携などを考慮して完全にCOBOLを脱却し、C/C++やJavaなどの言語でアプリケーションを再構築するというもの。2つめは、オンライン処理の部分を前述のような新しい言語で置き換えて、COBOLが得意とするバッチ処理の領域でのみ存続させるというもの。そして3つめが、COBOLならではの保守性の高さなどのメ

リットを生かすべく、継続利用するというものである。

とくに3つめのアプローチは、COBOL資産において確立されているビジネスロジックや機能を確実に継承していくという観点から有効であり、現在、多くの企業において採用されている。ただし、COBOLを継続利用するにせよ、その資産を新たな技術動向などに対応させていく、既存の資産を有効活用して業務プロセス改革を行うこと、いわゆる“モダナイズ”が重要である。

“モダナイズ”の必要性は新技術への対応ばかりではなく、法制度など社会動向の変化といったものにもモダナイズによって、システムをしっかりと追従させていく

### ①COBOLからの脱却

- 動的プログラミング言語（JavaScript、Ruby、Python、等）
- 静的型付き言語（Java、C、Swift、C++、等）

### ②オンライン処理はCOBOL脱却、バッチ処理はCOBOL継続

- オンライン処理は、業務端末要件も加味して、JavaScript、Ruby、Python、Java、C、Swift、C++、等でスクラッチ開発。
- バッチ処理は、大量データ処理を書きやすい（レコード入出力指向）、計算機は二進数でも人に合わせて十進数が中心（基本的に演算は十進数）のCOBOLを継続利用。

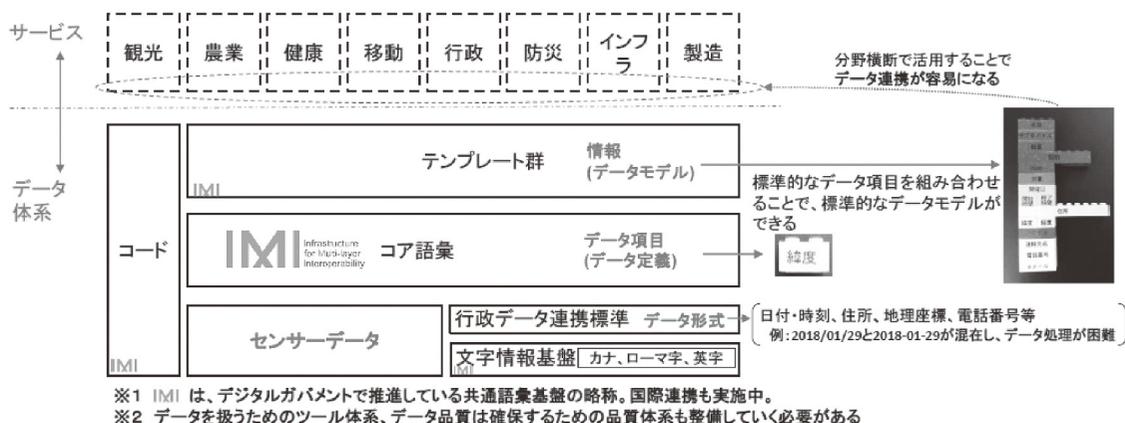
### ③COBOLの継続利用（保守性が高い）

- 静的言語（アプリケーションが稼働時に使うメモリ量を想定可能）
- 一つの文に多くの処理を詰め込まない（1ステップずつ処理が進み、理解しやすい）
- 長期にわたる互換性  
（国際規格が言語仕様を定義、互換性を保ちながら言語仕様が進化）

ターゲット

図1 システム再構築におけるCOBOL資産の取り扱い

## ■ 社会全体でデータ利活用するためには、基本データから積み上げた体系の中で相互運用性を確保していく必要がある。



## ■ データ体系は機動性の高い社会システムを作るための必須の基盤である。



※ 全体でコスト削減も実現

図2 データ標準の社会全体への展開。IT総合戦略室が推進するデータ体系（出典：2018年6月1日IMI意見交換会「データ流通社会へ向けた政府の取り組み」より抜粋）

必要がある。（図1）

そうした社会動向の変化の中でも、とりわけ大きなものとしてあげられるのが、現在、政府が進める Society5.0 実現に向けた取り組みである。IoT や AI、ビッグデータなどの技術を活用して、国内に抱える人口減少や高齢化、環境・エネルギー、防災といったさまざまな課題を解消し、社会そのもののあり方を変革していこうとするものである。

それに関連した一連の施策の中でも、既存システムへのインパクトという観点で、とりわけ注目すべきは、Society5.0 実現に向けたインフラ整備として、データの標準化と連携基盤の整備が推進されていることである。これは、政府が日欧米 10 億人規模のグローバルネットワークを想定し、そこで

のデータのやりとりにかかわる相互運用性を確保するために整備を進めている。

そうした流れの中、政府は Society5.0 実現に向けた「デジタル・ガバメント」の取り組みにおいて、「IMI (Infrastructure for Multilayer Interoperability: 共通語彙基盤) 共通データ標準としたデータ連携基盤」の構築を加速させている。これは、電子行政分野におけるオープンな利用環境の整備や、官民データ活用の推進の中核を担うもので、データに用いる用語や文字を共通化し、情報の共有や活用を円滑に行うための基盤である。

IMI は、大きく「共通語彙基盤」と「文字情報基盤」という2つのプロジェクトから成り、前者がデータとして用いられるさまざま

な用語の表記や意味、構造を統一して、分野横断的なデータの参照性を高めるためのものであるのに対し、後者は行政で用いられる人名漢字などを整備し、国際標準化を推進するものである。（図2）

とくに文字情報基盤については、戸籍に用いられてきた 55,270 文字の漢字と、住民基本台帳ネットワークの 19,563 文字の漢字を 58,842 文字に同定・整理して、国際規格に対応付けるというもので、外字を用いることなく、標準の文字セットのみで、誰もがそれらすべての文字を同一のコードで扱える世界をめざしたものである。

そのため、国内の標準化委員会をとおして、現行の Unicode の国際規格 (ISO/IEC 10646) に登録されていなかった文字を新たに登録し、さらに、字形差が小さ

くて統合して登録された文字の元の形を区別する一連のIVS (Ideographic Variation Sequence/Selector) を、Unicode コンソーシアムに対し提案し、国際規格が2017年12月に成立して、文字情報基盤の文字が使えるようになったことは感慨深い。これについては本誌2018年3月号の「文字情報技術促進協会主催シンポジウム顛末」にて袴田博之氏(日本電気)が寄稿している。

こうした動きを受けて問題となるのが、現在、Shift-JISをベースに8,000程度の文字を扱っているCOBOLプログラムを、いかに国際規格となった文字情報基盤で扱われる58,842文字を扱えるようにしていくかということである。新しい国際規格では、UTF-

8/16のベース文字コードに4バイトのIVSが指定されることになり、現行の2バイトを前提としたプログラムでは、文字のカウントや文字列の比較、コピー処理、あるいは文字データ領域の定義などについての全面的な見直しが必要になる。そのインパクトは、ほぼスクラッチからの開発に近い改修が必要となるほど大きなものである。

これに対し当社では、そうした問題を解消するためのソリューションを提供している。具体的には、COBOLの資産に手を入れることなく、「動的マッピング」と呼ばれる方法で実行時に文字データを適切にコード変換する仕組みを用意している。すでに行政分野をはじめとするユーザーが、このソリューションを活用して、国際規

格への対応、政府が整備した文字情報基盤に対応し、官公庁間、さらには官民を横断したデータ活用の推進に向けた準備を完了している。

Society5.0実現に向けた取り組みが加速する中、今後、企業が新たなビジネス価値を創出していくには、オープンデータの活用など、外部とのデータ連携が不可欠であり、既存のICT資産を最大限に活用しながら、そうした社会の動向に対応していくための“モダンイズソリューション”が注目されることに期待したい。(つづく)

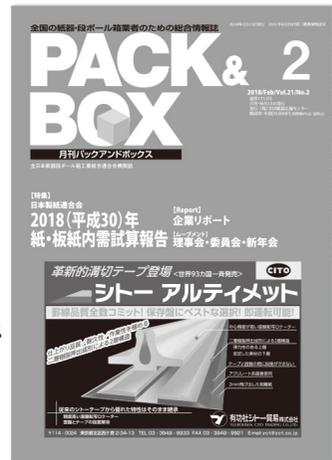
\*KAWAI, Takashi  
株式会社日立製作所  
公共システム事業部 全国公共システム  
第一本部 主任技師  
〒140-8512 東京都品川区南大井6-23-1  
takashi.kawai.rt@hitachi.com

## 全国の紙器・段ボール箱業者のための総合情報誌

# PACK & BOX

月刊バックアンドボックス

毎月15日  
発行



別冊も  
絶賛  
販売  
中

全紙器工連組合  
会社名鑑  
2017-2018 年度版

貼函図鑑  
貼函の世界

### 株式会社 全国紙器広報センター

〒130-0005 東京都墨田区東駒形 1-16-1 東京紙器センタービル 5F  
TEL.03-3624-9730 FAX.03-3624-9779  
mail:info@packandbox.co.jp http://www.packandbox.co.jp