EUC/EUDを前提とした文学研究システムのモデル化に関する考察 A Modeling for the Literature Studies System assuming EUC/EUD

富澤 浩樹^{1*}
Hiroki TOMIZAWA^{1*}

1 埼玉大学大学院文化科学研究科

Graduate School of Cultural Science, Saitama University

〒338-8570 埼玉県さいたま市桜区下大久保255

情報技術の進展と普及により、利用者がEUC/EUDを実践し、情報システムに関与できるシステム環境が整ってきた。しかし現在のところ、この観点からのシステムデザインはほとんどなく、人文系データベースシステムの議論の多くもその機能面からのアプローチが中心である。本論ではこの点を人文系データベースシステムが活性化しない一因としてとらえ、文学研究者のEUC/EUDの試みに着目したシステムデザインを試みた。

Recently, as EUC/EUD has become more common, the user is ready gradually to participate in an information system. But there is almost no system design from this viewpoint in a database system for literature studies. Furthermore, a system designer almost tends to discuss only the technical aspects of a computer. In this report, we regard this point as a cause which a system does not activate. And I attempt the system design surrounding literary research activities including EUC/EUD.

キーワード: EUC/EUD, 文学研究, 情報化社会, 情報システム分析

Keywords: EUC/EUD, Literary Studies, Information Society, Information Systems Analysis

1 はじめに

PC(Personal Computer)や携帯電話, 電子ブックリーダーなどの高性能な情報機器の普及を背景に, 情報技術に長けた専門家に頼らず, 利用者自らが情報システム[1]に関与する EUC(End User Computing)やEUD(End User Development)を実践することは, もはや日常となっている[2]。この社会環境の変化は, 人文系データベースシステム, 特に, 文学研究の基礎となる手稿や原稿などの一次資料の電子的な複製を, その利用者 (=研究者) に向けて公開する取り組み (本稿ではこのデータベースシステムを研究資料デ

ータベースと呼ぶことにする) に変化を与えている。顕著なところでは, テキストデータ中心から画像データ中心へ, スタンドアロン型からクライアントサーバー型への移行がある。しかし, その利用者の多様な要求を満たすことは難しく, 公開された研究資料データベースのメリットを享受していても, それぞれの視点から見た不満が常に聞こえてくる。著者の聞き取りによれば, たとえば, データの不備が訂正されない, システムマニュアルが作成されない, 欲しい資料のテキストデータがないなどがよく挙げられる[3]。その一方で, そのような利用者のほとんどが, 研究資料データベースの恩恵を享受

しながらも、EUC/EUD を実践することで自らその不足を補い、目的を達成している。

本稿では、そのような利用者の EUC/EUD の実践に着目する。すなわち、利用者をシステムの恩恵を享受するだけの受動的な存在ではなく、EUC/EUD を行い、積極的に貢献可能な存在としてとらえる。その上で、実際に EUC/EUD を実践した明星聖子教授(埼玉大学・ドイツ文学)の取り組みを対象にシステム分析を行う。そして、EUC/EUD の成果をどのように共有するかという観点から、文学研究システムのモデルについて考察する。

2 研究資料データベースと EUC/EUD

研究資料データベースの提供者は、公文書館、博物館、図書館などの公共施設や大学の附属施設、または研究者の研究グループ、そして出版社や企業などであり、成果の多くはインターネットを通して公開されている。近年では、利用者に向けてどのように提供すべきかという議論が散見されるようになってきたものの、提供者の関心の中心は、データベースの技術面・機能面にある[4]。これは、研究資料データベースの開発には、専門性と膨大な労力が必要であるためと考えられる。

しかし、たとえ利用者の要求を技術的に全て受け入れたとしても、いくつかの点で問題が生じる可能性が高い。まず、研究資料データベースのデータそのものの問題である。データベースの主要機能に検索や並べ替えがある。研究資料データベースにおいては、それら基本機能に影響を与える、転写(transcription)やタグ付け(markup)そのものが、研究対象となる場合も多い。データが直接的に基本機能に影響を与えているとなれば、様々な問題関心からシステムに接している利用者を抱えるシステムの信頼性を完全に担保することは難しい。一方で、システム開発における潜在的な問題がある。一般的に、

システム開発を巡っては、人間同士がなかなか分かりあえないという本質的な問題が関係しているため、要求定義は常にシステム開発の課題となっている[5]。つまり、提供者の取り組みだけでシステムを活性化させることは困難であると考えられる。

しかし、既存システムに不満を持っている利用者が、意識的にせよ無意識的にせよ EUC/EUD を実践し、既存システムの不足を補う役割を果たしていると考えれば、以上の問題が解消されないだろうか。利用者が主体的に情報システムに関与する EUC/EUD の意義は「データに意味を付与して情報から知恵をわき出させ行動に結びつけること[6]」ともいわれている。また、EUC/EUD においては、利用者自らが試行錯誤することで問題を克服するため、第三者が設計・開発を行うときに生じる要求定義の問題が発生しない。

そこで、次章より、EUC/EUD を前提とする文学研究システムの可能性について検討するために、明星のカフカ研究に焦点を当てる[7]。明星はフランツ・カフカの書く行為(エクリチュール)を解明する過程において EUC/EUD を実践しているためである。次章では、先ず紙メディア環境から EUC/EUD 環境に至るまでの明星のシステム環境について概観する。

3 対象のシステム環境

対象のシステム環境について、既に富澤と明星の先行研究[8]で示した現行システムの分析結果(図1)に基づき、①EUC/EUD に至るまでのシステム環境、②EUC 環境、③EUD 環境に分けて詳述する。なお、この図1[9]は、明星の文学研究をとりまくメディア環境を情報システム環境としてとらえ、その状況を示したものである。

3.1 第1～第3世代システム環境

図1 現行システムの分析結果(文献[8]より)

対象のカフカ研究では、研究者が参照する紙メディア上の研究資料＝「全集」から、EUC/EUDに至るまでに3つの世代を経ている。第1世代システム環境は、端的にいえば、プロット版カフカ全集を基盤とする研究環境である。第2世代システム環境は、1980年代以降の批判版カフカ全集に基づく研究環境である。そして、第3世代は、1990年代後半以降の写真版カフカ全集の研究環境である。

これら3世代は、どれも紙の書籍という情報メディア環境であるため、ひとつの世代とまとめてしまっているように見える。しかし、第1世代、第2世代と第3世代とでは、かたや活字かたやファクシミリという点で、媒体的な相違がある。さらに、第1世代と第2世代の間にも、前者は、一般読者を対象とした普及版、後者は、学術研究者を対象とした学術版だという点で相違がある。

なお、明星はこの点を、情報の信頼性に関わる問題として文学研究システム構想上のポイントであると、2010年3月に開かれた New Directions in Textual Scholarship 国際会議で指摘

している[10]。

3.2 第4世代システム環境(EUC)

第4世代以降の特徴は、PCがシステム環境に加わったことである。対象のカフカ研究においては、写真版カフカ全集に付録された、本の内容のPDFファイルを収めたCD-ROMを、読者＝利用者がPCを使って参照する。

ここで、新たな分析視点として、データの流に注目したDFD(Data Flow Diagram)を加え、PCを含むシステム環境を分析する[11]。その結果が図2である。図2の右上にある四角5は、PCを始めとする情報機器を示しており、利用者が情報機器を介しなければデータベースにアクセスできない点で、それまでの世代と明らかに一線を画していることが明示された。

この点について、2010年3月に開かれた国際会議で明星は、第4世代システム環境を、いわゆる一般的な電子書籍の環境と繋げる議論を展開した。そして、その特徴について、コンテンツとそのコンテンツを載せる搬送体が切り離された点をあげている[12]。この明星の指摘は、対象のカフカ研究だけではなく、広く一般にも

図2 第4世代/第5世代システム環境(現行物理DFD)

EUC が浸透していることを示しており、そのためのシステムデザインが必要であるとも考えられるが、本稿ではそれを示唆するに留めたい。

3.3 第5世代システム環境(EUD)

図2右上の四角5で、提供されたデータを指定されたツールを用いて閲覧するばかりでなく、独自に注釈や解釈データを作成し、それを系統的に整理するためのソフトウェアを利用者が独自に開発すれば、第5世代システム環境となる。明星は、第4世代システム環境として提供されたデジタル資料が紙資料特有の制約を継承していることに不満を感じ、①データのグルーピングの固定性、②データのシークエンスの固定性の打破を目指し、プログラマの菅生紳一郎の協力を得て研究ツール LiterarySpace を開発している[13]。

この LiterarySpace の開発思想については、2008年6月にフィンランドのオウルで開かれた Digital Humanities 国際会議での明星らの発表に詳しい[14]。それによると、LiterarySpace は、画像データとテキストデータとコントロールデータの3種類のデータを3次的に組み合わせたデータ構造からなる。画像データの1次的な蓄積は、「経験」の層と名付けられている。この3次的なデータ構造の全体こそが、まさしく知識創出の構造のモデルであるという。そして、

その画像データひとつひとつに、1 対多対応で、2 次元的に、transcription や interpretation や translation の結果としてのテキストデータが関連づけられる。この 2 次元的な層は、「解釈」の層と呼ばれている。そして、その画像データとテキストデータのセットを、さらに、自由に組み合わせ、シークエンスを入れ替えて、3 次元的な空間に、新たなデータファイルを創出させた層が、「創作」の層といわれる。

本稿の問題意識に立ち返れば、EUC/EUD の成果といえるこのデータファイルを、いかに位置付け、どのように活用するかが、文学研究システムをデザインする上での要求項目の一つといえる。

4 文学研究システムのデザイン

4.1 現行システム環境の機能分析

ここで、図1の現行システムに現れた複数の機能について分析を行う。具体的には、同じ区分に入れることのできる機能について、より抽象的な置き換えを行った[15]。図1で現れた機能は、「管理・公表」、「編集・公表」、「参照」、「入力」、「編集」である。これら機能が何を行っているのかに注目してまとめると、表4で示されるような5つの領域として区分できる。すなわち、現物資料のアーカイブを行う機能(領域0)、:画像データとテキストデータを提供する機能(領域I)、テキストデータを

表1 現行システム(図1)の機能分析

現行システムの機能(図1より)	(関連)	区分できる領域
管理・公表	→	領域0: 現物資料のアーカイブを行う
編集・公表	→	領域Ⅰ: 画像データとテキストデータを提供する
参照	→	領域Ⅱ: テキストデータを参照する
入力	→	領域Ⅲ: 画像データを参照する
編集	→	領域Ⅳ: 画像データとテキストデータを活用する

参照する機能（領域Ⅱ），画像データを参照する機能（領域Ⅲ），画像データとテキストデータを活用する機能（領域Ⅳ）を持つ5つのサブシステムの総体として，対象の文学研究が成立していると考えられるのである。

4.2 EUC/EUD を前提とした文学研究システムのモデル案

第5世代システム環境では、デジタルデータを自らのPCで参照するだけでなく、それを自らのコンピュータに取り込み、自ら編集し、そして新たな解釈を付加する機能を担っていた。したがって、EUD環境が次に求めるのは、新たな個人的エディションを作り出し、それを公表したり、既存のデータベースシステムに研究成果を反映させるためのシステム環境であるともいえる。そこで、表1で示した領域に、研究成果を反映させるための機能を追加し、EUC/EUDを前提とした文学研究システムのモデル化を試みた。その検討結果が図3である。なお、モデル化の際には、SSADMの手順に則り、誰が、どのように、予算などの現実的な実現可能性については一先ず棚上げしている。

先ず、唯一無二の現物から領域0へ矢印を引
 き、そこから画像化およびテキスト化をした上

でデータベースに保存する領域Ⅰに繋げた。ここで、テキストと画像は構造化され、関連付けられることによって公開可能となる。このときのインターフェースを領域Ⅱ(テキスト)および領域Ⅲ(画像)として並列に繋いだ。そして、領域Ⅳ(EUC/EUD)ではそれらからデータを取得し、研究成果としてのデータセットを作成して出力する。しかし、それを受け入れるシステム領域が存在しないため、新たに「研究成果データベース」を置いた。領域Ⅰのデータベースはこの機能を経て更新される。

5 問題提起(おわりに代えて)

本稿では、EUC/EUD によって作成されたデータを、研究資料データベースに反映させるための抽象的なモデルについて、対象のカフカ研究に基づいて検討した。図3を一目見て問題と指摘できるのは、たとえ各領域のシステムが分散的に存在することになったとしても、どの現物を基に画像データやテキストデータが作成されたか同定する必要があるということである。この問題について、明星は現物の名付けの問題として指摘している[16]。システムデザインの視点からみても、それが決まらなければ構想もままならないことが明らかである。先ずは文学研究者と現物資料管理者との今後の議論に期待したい。

また一方で、2010年3月に開かれた国際会議において、デジタル環境における公表の概念の変容と文学研究活動との関連の重要性についても明星により指摘された[17]。すなわち、EUC/EUD環境で作成される画像データに対する解釈の結果であるテキストデータについて、どの段階でどのように公表するのか、またどのような規則で関連付けるのかなどについて、文学研究者間での合意形成が必要であるという。この指摘は、文学研究システムをデザインするとき、画像データに

図3 対象の文学研究システムのモデル案

対するテキストデータの関連を、システム全体でどのように保つかという大きな要求項目となる。この要求に対しては、図3の各領域で稼働する機能的システムのハードな課題としても、利用者組織の合意形成などのソフトな課題としてもとらえられよう。また、公表の問題については、補助線(点線)を図3に引いて用いれば、システムデザインの視点からの議論が可能である。たとえば、明星のLiterarySpaceは、境界Ⅰの私的領域(領域ⅠからⅣ)をカバーする研究ツールといえる。また研究関心によっては、境界Ⅱを検討する可能性もあろう。この場合、テキストや画像の新たな参照の仕方そのものが研究対象となる。そして、現在稼働中の多くの文学資料データベースは、境界Ⅲの公的領域をカバーしているようにみえる。一般化をおそれずに考察すれば、利用者の声によっては、一部の機能領域をサブシステムとして私的領域に公開することで、EUC/EUDを取り込み、システム環境を活性化する可能性があるといえよう。

かつて、企業システムにおいてEUC/EUD環境を活性化するためには、情報システム部門と利用者部門との協力が不可欠といわれた[18]。文学研究システムにおいても、共通の現物を扱う研究者や現物資料管理者、そしてシステムデザイナーとの連携が欠かせない。今後も文学研究者らの協力を得ながら、システムデザインの視点から議論を深化させていきたい。本稿が文学研究システムに関する今後の議論に貢献できれば幸いである。

注・参考文献

- [1] 本稿では情報システムをコンピュータシステムと同義とする一般的な理解ではなく、以下の文献で示された「組織体および個人の活動に必要な情報の収集・蓄積・処理・伝達・利用にかかわる仕組み」とする広義にとらえる。浦昭二他:『情報システム学へのいざないー人間活動と情報技術の調和を求めて』培風館, 1998
- [2] 田村幸子:「エンドユーザコンピューティングと情報リテラシー」,『九州産業大学商経論叢』第37号, pp. 95-109, 1996
- [3] ある研究者が、文献上の差異と思われていたことが実は研究資料データベース上の登録ミスであったことを発見した。この研究者はこの事実を管理者に指摘したが、修正に応じてもらえなかった。現在もそのままであるという。
- [4] 以下の文献で師茂樹は、「開発された結果ばかりに目がいてしまい、モデルや方法論的な検討はあまり重視されないようである」と、研究資料データベースの研究状況について所感を述べている。高田時雄他:『人文情報学シンポジウムーキャラクター・データベース・相互行為ー報告書』, 京都大学21世紀COEプログラム「東アジア世界の人文情報学研究拠点 漢字文化の全き継承と発展のために」, 2007
- [5] Checkland, P. and Scholes, J., *Soft Systems Methodology in Action*, Wiley, 1990. (妹尾堅一郎監訳:『ソフトシステムズ方法論』, 有斐閣, 1994)
- [6] 佐藤修他:『エンドユーザコンピューティング』, 日科技連出版社, 1996
- [7] 明星の具体的な研究活動については、以下の文献を参照。明星聖子:『新しいカフカー「編集」が変えるテキスト』, 慶応義塾大学出版会, 2002
- [8] 富澤浩樹, 明星聖子:「文学研究活動に着目したEUC/EUDに関する考察」, 経営情報学会 2009年秋季全国研究発表大会, pp. 288-291, 2009
- [9] 四角形は、対象のシステム環境に存在する利用者(U: User), 提供者(S: Supplier), PC(COM: Computer)の役割を示している。その右横に機能を、その下には機能の説明を記した。長方形はデータ庫(=全集)である。矢印は、四角形と長方形の関係を示す。詳細は注8の文献を参照のこと。
- [10] Kiyoko Myojo: "Public Editing and/or Private Editing: Changes in the Concept of Publication in the Digital Age", International Conference "New Directions in Textual Scholarship", Saitama, 2010/3/26
- [11] Geoff Cutts: "Structured Systems Analysis and Design Methodology 2nd.ed", Alfred Waller Limited Publishers, 1991 (浦昭二監訳:『情報システムの分析と設計ーSSADMとその実践』, 培風館, 1995)
- [12] 明星は、その国際会議(注10)で、従来の紙の書籍は情報とその情報の搬送体が一体化したものだだったが、第4世代システム環境ではそれらが分離しうると指摘した。その上で、近年の電子書籍の普及をめぐる社会的な大きな進展は、人類の書物の世界が全体として大きく第4世代環境に進んでいるという旨を、英語で述べている。
- [13] 研究ツールLiterarySpaceへの要求の詳細については、注8の文献を参照のこと。
- [14] Kiyoko Myojo; Shin'ichiro Sugo: "LiterarySpace: A New Integrated Database System for Humanities Research", *Digital Humanities* 2008, pp. 260-262, 2008
- [15] システムデザインでは論理化といわれる(注11参照)。
- [16] 明星聖子:「文学研究資料の未来をめぐる一考察ー編集文献学とアーカイブズ学の狭間でー」,『アーカイブズ学研究』第11号, pp. 92-111, 2009
- [17] 注10参照。
- [18] 注6参照。