

集合知的手法による学際的学問分野の知識体系構築システムの プロトタイピング

莊司 慶行[†] 伊藤 一成[‡] 増永 良文[‡]

[†] 青山学院大学社会情報学研究科 〒229-8558 神奈川県相模原市淵野辺 5-10-1 O-206

[‡] 青山学院大学社会情報学部 〒229-8558 神奈川県相模原市淵野辺 5-10-1

E-mail: [†] c8108008@aoyama.jp, [‡] {kaz, masunaga}@si.aoyama.ac.jp

あらまし 現代社会における諸問題は複雑性を増しており、理系文系というように切り分けられた既存の学問分野による解決が困難になっている。そこで昨今注目を集めているのが学際的アプローチであるが、領域横断的な学問は、その学問分野を明確に定義し、表現することが困難である。本稿では、青山学院大学における社会情報学の定義の策定をモデルとし、Wikipediaに見られるような集合知形成手法を用いた学問体系の策定を可能にする環境構築を行う。社会情報学の学問体系 SIBOK (Social Informatics Body Of Knowledge)を共同的かつ動的に作成するため、学問体系の記述形式である BOK+を提案し、多人数で学問体系の編集できる環境のプロトタイプとして Semantic MediaWiki をベースとした知識体系構築システム“BOK コンストラクタ”を作成した。

キーワード SIBOK, BOK, Body of knowledge, Semantic Mediawiki, wiki, 社会情報学, 学問体系, 集合知

Prototyping of a Construction System of a Body of Knowledge of an Interdisciplinary Field of Study based on a Collective Intelligence Approach

Yoshiyuki SHOJI[†] Kazunari ITO[‡] and Yoshifumi MASUNAGA[‡]

[†] Aoyama Gakuin University Graduate School of Social Informatics 5-10-1 O-206 Fuchinobe, Sagamihara city,
Kanagawa 229-8558 Japan

[‡] Aoyama Gakuin University School of Social Informatics 5-10-1 Fuchinobe, Sagamihara city, Kanagawa 229-8558
Japan

E-mail: [†] c8108008@aoyama.jp, [‡] {kaz, masunaga}@si.aoyama.ac.jp

Abstract Today, The solution of the academic problem by conventional study field that divided into the science course and the liberal arts course becomes difficult, because contemporary society is too complex. Therefore, an interdisciplinary approach becomes more important. However, it is very difficult in an interdisciplinary study area to express and to define its study field. In this text, we construct the environment that enables the decision of the study field with collective intelligence approach shown in Wikipedia, through the decision of the definition of social informatics in Aoyama Gakuin University as a model case. To make the body of knowledge of social informatics (SIBOK: Social Informatics Body Of Knowledge), we advocated BOK+ that was the description form of the study system as a collaborative and dynamic document. And, we implemented the body of knowledge construction system based on Wiki named "BOK constructor" as the prototype of the environment to be able to edit the study system by the large number of people collaboratively.

Keywords SIBOK, Social Informatics, BOK, Body of knowledge, Semantic Mediawiki, wiki

1. はじめに

1.1. 研究背景

近年、複雑化した社会情勢を背景に、既存の学問分野に当てはまらない高度な問題を解決する必要性が増してきている。そのため、複数の学問分野を横断し研究を行う、学際的なアプローチが注目を集めており、新しい学際的な研究領域を対象とする学部や学科が著し

く増加している。一例として、青山学院大学では、こういった難問の解決にあたるため、2008年度に社会情報学部を新設した。社会情報学の研究機関として、札幌学院大学、群馬大学、大妻女子大学、青山学院大学に社会情報学部が、また京都大学大学院、東京大学大学院に社会情報学専攻やコースが設置されている。しかしながら、社会情報学とは極めて若い学問分野であ

り、現段階では、社会情報学とはどのような学問分野であるかを明瞭に表す定義も、また学問分野に対しての研究者達による合意も存在しない。社会情報学とは何かを記述しようとした一例として、Wikipediaを参照してみると、以下のように記されている。“Social informatics refers to the body of research and study that examines social aspects of computerization [1],”しかしながら、この記述は社会情報学の定義を説明するにあたって情報量に乏しく、また不明瞭であるといわざるを得ない。社会情報学とはどのような学問で、どのような問題系を対象とするのかを明確に記述した学問体系 (Body of knowledge) が早急に必要である。しかし、社会情報学は情報科学と社会科学の融合した学際的な学問分野であり、文系、理系という従来の学問分野の垣根を越えた研究者が所属している。そのため、どちらかの領域に属するような、特定の個人による学問分野の定義は困難である。これまでの研究で、このような環境下で適切な学問体系を構築するための手法として、ボトムアップアプローチで多分野多人数の研究者の意見をくみ上げ、集合知形成手法を用いて共同的に学問体系を作り上げる必要があり、またその道具としてコンピュータネットワークならびにWebアプリケーションが有用であるとの認識に至った[2]。

当研究では、青山学院大学における社会情報学の定義決定を例にとり、共同的で動的なドキュメントとしての社会情報学の学問体系 SIBOK (Social Informatics Body Of Knowledge)の作成を目的とし、そのために必要な学問体系の記述形式であるBOK+を提案し、多人数で学問体系の編集可能な環境のプロトタイプとして Semantic MediaWikiをベースとした知識体系構築システムを作成した。

1.2. リファレンスモデル

学問分野の定義を明確に記述することは、複数の学問が融合した学際領域において、とりわけ重要性が高い[3]。本稿で例として挙げ、知識体系記述の対象とした BOK は、青山学院大学における社会情報学という極めて狭い領域をスコープとしていた。このスコープを二つの方向性で拡大することで、以下のように学問分野を相互参照し、学問領域どうしを相対化できるようになると考えられる。

- (a) 学問全体をスコープとし、社会科学と情報科学の融合 BOK としての社会情報学 BOK を構築する。社会情報学のみに含まれる部分、社会情報学と社会科学、情報科学のどちらかのみと共有する部分、社会情報学と社会科学と情報科学のすべてに共通する要素を明らかにすることで、社会情報学がどのような学問かを相対的に明らかにする。

- (b) 社会情報学をスコープとし、青山学院大学の掲げる社会情報学と他の大学の掲げる社会情報の融合 BOK として、社会情報学全体の BOK を表現する。いくつもの異なった団体による社会情報学の BOK を融合させ、それぞれを比較することで、各団体の特色を差別化する。

以上のように、明確に記述された学問分野の知識体系は、参照モデル (リファレンスモデル) として活用できる。これにより、学問分野同士の相対化が進むほか、より活発な領域横断型のコラボレーションを生み出す呼び水となることが予想される。

1.3. 論文構成

本稿における以降の構成について述べる。2 章では、学問分野の知識体系をいかに記述するかについて述べ、学問分野記述のための構造である BOK+ を提案する。3 章では BOK の共同構築をするために必要な環境と、その実装のプロトタイプについて述べ、4 章でまとめと今後の課題について述べる。

2. 学問分野の表現

2.1. BOK

ある学問分野の定義を記述する際、学問体系 (Body of knowledge, BOK) がよく用いられる。著名な例として CSBOK[4]がある。CSBOK は IEEE-CS ならびに ACM による共同タスクグループによって作成され、公開されている。CSBOK ではコンピュータサイエンスに関する学問の定義を、文書における見出しのように 3 段階に階層分けされた木構造グラフと、それを説明する短い文によって表現している。

この木構造グラフにおいて、最高位のレイヤをエリアと呼ぶ。エリアは学問分野内のサブフィールドであり、複数のユニットによって構成される。例として、情報科学 (Computer science) という学問分野において、“Discrete Structure” はエリアである。

ユニットは、当該のエリア内での、個別のテーマごとにまとめられたモジュールであり、複数のトピックによって構成される。例として、“Functions, relations, and sets”, “Basic Logic” などのユニットがまとまって、“Discrete Structure” というエリアを構成する。

トピックは、三階層の木構造のうち最下位のレイヤであり、その学問分野に置いて実際に対象となっている個々の論題や単語である。例として、“Functions, relations, and sets” というユニットを構成するトピックは、“functions”, “relations”, “sets” などである。

BOK の他の例として、PMBOK や SWEBOK などがあるが、これらの学問体系の記述は、このような木構造で表される見出し構造を持つ文書であったり、また自然言語によって書かれた文章である。こういった BOK は、編集をするにあたって会議などで割り振りを

行い、担当区分を決めた上でトップダウンに編集する必要がある。また共同編集するにあたって情報量が乏しく、専門の異なった研究者の多数いる学際的な学問分野において、領域をまたがった共同編集を行う事が難しい。また、既存の BOK は静的な完成物を作成することを前提としており、編集をかさねることによって生じる、異なる領域間でのコラボレーションが発生せず、学問体系同士が混ざり合わない寄せ集めのような BOK になってしまうことが考えられる。

2.2. 共同的な BOK 構築

一般的な BOK の構築は、CSBOK の例を典型として、大規模なタスクグループを作成し、領域を分割したうえで、作業班ごとに専門家が各々の担当区分についてトップダウンに記述するという方法がとられる。しかしながら、新規領域の BOK を策定する際には、最初に固定的な領域の分割を行い作業班を設けることが困難である。また、社会情報学のような学際領域においては、特定の分野に所属する一部の研究者が作業班を決定することは、完成される BOK が必要十分に欠く、恣意的なものになりかねない。そこで、我々は、Wikipedia に見られるような集合知形成手法を用いた、ボトムアップアプローチに基づく BOK 構築環境を実現する。この環境下では、学問分野に所属する全ての研究者が自由に BOK の編集に参加し、各々の専門分野について記述するばかりでなく、他の研究者の専門分野と相対化し、関係性を記述できる。

このような編集を可能にする上では、個々の研究者が自身の対象とする学問の構成要素について構造が何もない状態から記述し、他の研究者がその記述に対して十分な情報と論拠を読み解いて自身の記述と構造的に繋げることが可能でなければならない。そのために、研究者は、まず自身の研究領域における論文や、実際に用いられている教材を提示し、それらを十分な説明が伴った学問の構成要素とする必要がある。すなわち、何もない状況から共同的に BOK を構築するためには、

- (1) 論拠となる論文や教材を提示する
- (2) 論文や教材を分割し、ある学問の構成要素を説明する文書を複数作成する
- (3) 学問の構成要素同士の関係性を記述する

というプロセスをたどって、BOK の編集に参加する。一方、ある程度学問体系の記述が進み、BOK の構造を補強する場合には、異なったアプローチが必要となる。すなわち、

- (1) 学問体系として足りない構成要素を削除追加したり、構成要素同士を分割マージしたり、構成要素同士の関係性を編集する
- (2) 変更した学問の構成要素について、具体的な説明を加える

(3) 説明の論拠となる論文や教材を提示するというプロセスによって、BOK の記述を改善する。これら二つの編集形態を組み合わせ、多人数によって BOK の編集を重ねてゆくことで、堅牢な学問体系の記述が可能になると考えられる。

さて、ここで作成される学問体系の記述は、先に集合知の例として挙げた Wikipedia のような、個々の記事の集合体であり互いの関係性に乏しい百科事典形式ではなく、十分に構造化された文書である必要がある。すなわち、学問分野として、学問の構成要素同士の構造的な関係性、学問の構成要素の具体的な説明や、その論拠となる論文や教材について、構造的に編集できなければならない。これらをふまえ、学際領域の学問分野を構築するため、共同的で動的な編集が可能な学問体系記述形式として、BOK+構造を提案する。

2.3. BOK+構造の提案

社会情報学という、社会科学と情報科学の融合した、極めて若い学際の研究領域の定義を十分に記述するため、共同的で動的な文書形態として BOK+を提案する。BOK+の構造の概要を図 1 に示す。BOK+は 3 つの領域から構築される。領域内のそれぞれのインスタンスは意味リンクによって互いにリンクされている。

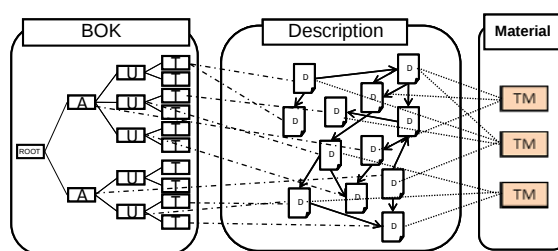


図 1 BOK+の構造

Fig 1. Outline picture of BOK+ structure.

図 1 の最左の領域が BOK である。これは CSBOK などにおける BOK と同様に、木構造で学問体系を表現したものである。図中の木構造グラフにおける一つ一つのノードを、BOK エレメントと呼ぶ。BOK エレメントはそれぞれが学問の構成要素である。それぞれの BOK エレメントクラスのインスタンスは、親ノードと子ノード、自身の名前の他に、同義語と同型異義語へのリンクを持つ。これは学際領域において、同一の概念が全く別の分野で用いられる場合や、異なった概念が同一の名称で呼称される場合があるためである。複数の学問分野の BOK を融合した、ある学際領域の学問体系を表現する際、BOK における木構造の制約が枷になる場合がある。このような場合、図 2 のように、異なった学問分野における同一の概念に対して同義語

のリンクを張ることで、一つのノードが二つの親を持つと同じ構造を表現できる。また、同型異義語の例として“ネットワーク”の語が挙げられる。“ネットワーク”は情報科学という学問分野ではコンピュータの構成する情報網を指し、WANやLANなどを指す。一方で、社会科学の領域に置いては、“social network”を指す。このような際に、同一の見出し語を持ちながら、使われる文脈や領域によって指し示す概念が大きく異なる語の持つ関係性を適切に表現するため、これらの単語間を同型異義語のリンクで結ぶ。各ノードはそれぞれCSBOKに倣ってエリア、ユニット、トピックという階層情報を持っている。BOK+では複数のBOKを融合させる事も念頭に置いているため木構造の階数を3段階に限定しないが、表示の際にこの階層情報に従って3階層に変形した上で表示することが可能となる。

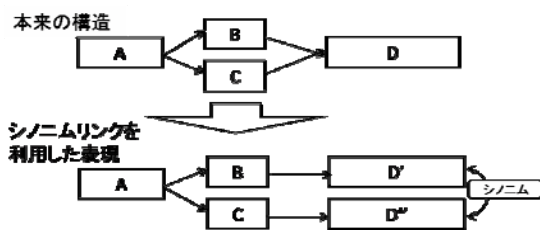


図2 類義語を利用した親を複数持つ要素の表現
Fig 2. Description of a node with more than one parents using a “synonym link.”

図1における中央の領域は、学問の構成要素を説明する文書である description である。一つ一つのBOKエレメントは、それぞれが学問の構成要素を表しているが、その構成要素が実際にどのようなものなのかの記述を持たない。そのため、学問の構成要素を説明する自然言語で書かれた description と紐付けされる。description は、説明するノードや論拠となる教材、関係ある他の description などといった構造情報の他に、自然言語で書かれた本文をメンバとして持つ。この本文中で他の description を参照したり、明示的に他の description との関係を明示することができる。これによって、description 群は、BOK エレメントほど作為的ではないが、潜在的にセマンティックネットワークのような構造を持つ。

図1の最右の領域は、Teaching Material である。Teaching Material の一つ一つのノードは、実際の論文や教材の実体を表す。これらの論文や教材は、description の論拠となる。Teaching material の各インスタンスは、シリアルナンバによって管理され、実際のファイルへのポインタ、ファイルの概要、ファイル名をメンバとしてもつ。

互いに紐付けされた、これら3種類のクラスのイン

スタンスによって、学問体系を共同的に記述出来ると考える。

3. BOK 共同構築環境の実現のためのプロトタイプ

3.1. SMW の利用

BOK+を共同で編集するためのプラットフォームとして、MediaWiki に注目した。MediaWiki は Wiki (すなわち WikiWikiWeb) クローンの一つであり、Wikipedia に採用されていることで知名度が高い。Wikipedia の成功から見て取れるように、集合知形成の道具として Wiki は極めて秀でた機能を有している。標準で、個別の記事に対しての編集機能、ユーザマネジメント機能、履歴管理機能が備わっている。

本研究では、構造化された文書を MediaWiki 上で取り扱うため、Semantic MediaWiki という MediaWiki 用のエクステンションを適用した。Semantic MediaWiki は SMW project によりオープンソースで開発されており、MediaWiki にタグや記事同士の関係など、意味情報を付与できるようにする。以降、便宜上、Semantic MediaWiki エクステンションを適用した MediaWiki のことを Semantic MediaWiki と呼ぶ。Semantic MediaWiki では、記事同士の関係性を、記事中の Wiki ワードリンクにキーワードを埋め込むことによって表現する。例として、“イクラはシャケに対して Is_child_of の関係である”という事実を記述する場合には、“イクラ”という記事内に、“[[Is_child_of::シャケ]]”と記載すればよい。このように記述された意味情報は、Semantic MediaWiki 上で意味情報としてナビゲーションに利用できるほか、RDF 形式でエクスポートし、様々な用途で利用することが出来る。

Wiki は極めてユーザフレンドリであり、また Wikipedia の知名度も手伝って、社会情報学という非情報系の研究者が多い環境においても、コンピュータの専門家でない研究者にとっても違和感なく利用することが出来る。また、Semantic MediaWiki をプラットフォームとして用いることで、既存の Wiki 向けのマイニング手法を適用可能である他、MediaWiki 用の新しいエクステンションを適用できるなど、副次的な利便性を得られた。

3.2. BOK コンストラクタ

Semantic MediaWiki を利用して学問分野を共同編集する環境を作成するため、BOK コンストラクタという Web アプリケーションを作成した。

BOK コンストラクタのシステム概要図を図3に示す。

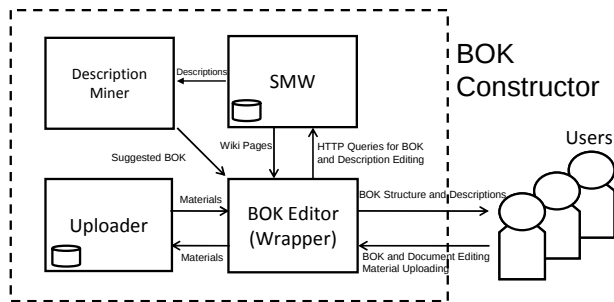


図 3 BOK コンストラクタの構成
Fig 3. Diagram of BOK Constructor.

BOK コンストラクタを作成するにあたり，MediaWiki 内部に手をを入れて改造したり，MediaWiki エクステンションとして実装したりせずに，MediaWiki をシステムの一部として取り込んだシステムとした．BOK コンストラクタの各モジュールは MediaWiki のフロントエンドのように振る舞い，MediaWiki は BOK コンストラクタにおいて，データベースのように利用される．

BOK コンストラクタは Semantic MediaWiki を合わせて 4 つのモジュールから成り立っている．

BOK エディタモジュールは BOK コンストラクタ全体のフロントエンドであり，ユーザは全てのモジュールに BOK エディタ経由でアクセスする．Semantic MediaWiki 上に保存された構造化された文書を編集する場合に，BOK エディタは Semantic MediaWiki 内の複数の記事を読み出し，木構造として可視化しユーザに提示する．ユーザはグラフデータを可視化されたまま編集することができる．BOK エディタを通して Semantic MediaWiki にアクセスすることにより，社会情報学という学際領域において，セマンティック Web や WikiML に明るくない教員と，情報系の教員が，同じ土俵に立って編集に参加できる．

アップローダは教材や論文をアップロードするためのモジュールである．アップローダへのアクセスは BOK エディタを通して行われる．BOK エディタはアップロードされた全ての教材に対してシリアルナンバを付与し，Semantic MediaWiki 上に TM 型の記事を一つのファイルに対し一つ生成する．これにより，Semantic MediaWiki 上で仮想的にファイルの実体に対して意味リンクを貼ることが出来る．

ディスクリプションマイナモジュールは，ユーザによって教材や論文をもとに作成された Description に対しマイニングを行い，ユーザが BOK を編集する手助けをするためのモジュールである．Description について自然言語処理やリンクの解析を行い，その Description が BOK のどこに紐付けされるべきか，また BOK に新しいノードを追加すべきかを推薦する．

3.3. BOK エディタ

BOK エディタは，BOK コンストラクタ全体のフロントエンドであり，ユーザは Semantic MediaWiki 内のデータを編集する場合や教材をアップロードする場合，構造化された情報を可視化して表示する場合など，BOK エディタを経由して他の全てのモジュールにアクセスする．

BOK+の各クラスを Semantic MediaWiki 上で表現するために，BOK エディタは構造情報を WikiML で表現し，Semantic MediaWiki に書き込む．各クラスはネームスペースで分離され，実際に記述される WikiML はクラスによって異なる．例として，BOK エlement 型のノードの実際のメタデータの記述方式を図 4 に示す．図 4 の BOK エlement は，“Relational database design elective”という学問の構成要素が”Information Management (IM)”という要素の下位概念であることを表し，さらに下位の概念として“Database design, Functional dependency”などがあることを表している．また，このノードを説明する description はなく，このノードはユニットであることなどを表している．

```

{{element
|name=[[Element:Relational_databases_elective]]
|parent=[[Hyponym::Element:Information_Management(IM)]]
|child=[[Hyponym::Element:Mapping_conceptual_schema_to_a_relational_schema]]
/[[Hyponym::Element:Entity_and_referential_integrity]]
/[[Hyponym::Element:Relational_algebra_and_relational_calculus]]
|class=[[is_a::UNIT]]
|heteronym=
|homonym=
|document=}}

```

図 4 WikiML によるエレメントクラスの表現
Fig 4. Sample of WikiML.

3.4. プロトタイプピング

BOK コンストラクタを実際に使用して BOK+構造を編集するプロトタイプとして，Web アプリケーションを実装した．このアプリケーションは，青山学院大学社会情報学研究科において，文理両方の教員が使えるよう，必要最低限の簡潔なインタフェースとして固定的な 3 つのビューを持つ．3 つのビューはそれぞれ，BOK の木構造の編集と description や Teaching Material のアップロードを行う統合的ビューであるメインビュー，複数の description と Teaching Material 間の紐付けを一括で編集できる description - TM ビュー，そして BOK 内の類義語と同形異義語の紐付けを一括編集できる element - element ビューである．メインビューのスクリーンショットを図 5 に表す．この Web アプリケーションでは，3 つのビューを横断的に利用し，木構

造の BOK 編集, description の追加編集, Teaching Materials のアップロードおよびダウンロード, そしてすべての要素間の紐付とその変更が出来る。

実際に BOK+構造の編集が十分行えるかの試験として, 各クラスにおける全てのオペレーションについて確認を行った。例として, element クラスにおいては, 他のノードの子へ自身を移動させる機能, あるノードを親とする新規ノードの作成する機能があれば必要な全てのオペレーションが行える。この際, ノードの削除は, 削除したいノードをゴミ箱ノードの子へ移動することで対応した。また, ノードの移動に関して, グラフ上に円環状の部分ができるような移動については, 禁則処理を適用し, 適切にエラーを出力し実際にオペレーションを行わないという制約を課した。これらの操作を Semantic MediaWiki 上で表現するため, 依存関係にある全てのノードの WikiML について書き換えを行い, また読み出した上で GUI 環境下で編集できることを確認した。

BOK+構造を実際使用して学問体系を記述でき, また異なった異種 BOK を適切に融合出来るかの検証実験として, BOK コンストラクタに試験的に CC2001 と Library of Congress Classification [5]の一部について, BOK エディタを通して入力した。入力されたデータは CC2001 よりすべてのエリア, Information management (ja)以下のすべてのユニットとトピック, Social and Professional Issues (SP)以下のすべてのユニットとトピックであり, また LCC より CLASS H のサブクラスをすべてと, "Industries. Landuse. Labor", "Finance", "Social pathology. Social and public welfare. Criminology"のサブクラスである。

実装されたインタフェースを利用し, ここで入力されたデータを BOK 構築用のシードコンテンツとして用いることで, 多人数での BOK+構造の共同構築が行える。

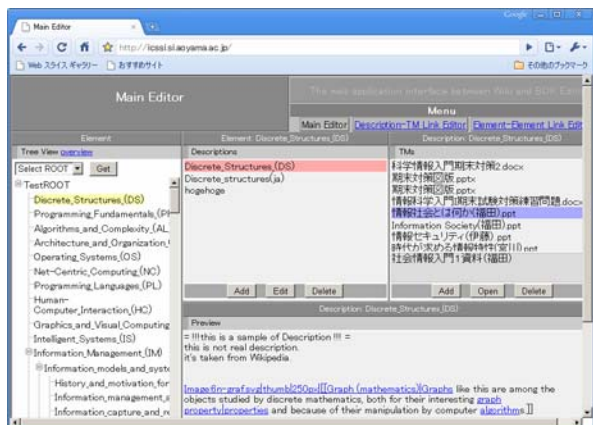


図 5 Web アプリケーションのスクリーンショット

Fig 5. Screen shot.

4. 結論と今後の課題

学際的な学問分野における集合知的手法を用いた知識体系の策定を実現するため, その第一段階として集合知形成手法で必要十分な学問体系を記述するための文書形態としての BOK+, そして BOK+を健全に共同構築するための BOK コンストラクタについてそれぞれ提案し, またプロトタイピングを行った。学際領域における学問体系の明確な定義は学問および研究機関を比較相対化するため極めて重要であり, その共同構築のためには BOK+のような動的かつ共同的文書の形態が必要不可欠である。

本稿における研究では, 学問分野の定義を共同的に策定するために, システムとしての BOK+構造および BOK コンストラクタについて研究を行ったが, 優れた学問体系を記述するためには, システムの話だけに留まらず, 参加者間における編集のルールや, 活発なコミットメントを促進する為の仕組み, 健全な編集を行うのに必要なコミュニティの形態等についても研究が必要である。また, BOK コンストラクタにおける UI のユーザビリティ, BOK+を閲覧する際に人間にとって見やすい表示方法, 同時多人数編集への適切な対応についても研究が必要であると考えられ, 今後の課題として今後研究を行う。

参 考 文 献

- [1] http://en.wikipedia.org/wiki/Social_informatics
- [2] Yoshifumi Masunaga, Yoshiyuki Shoji, Kazunari Ito: Collective Intelligence Approach for Formulating a BOK of Social Informatics, an Interdisciplinary Field of Study, Proceedings of the 5th International Symposium on Wikis (WikiSym 2009), Article No.: 34, Orlando, FL, USA (2009).
- [3] 増永良文, 莊司慶行, 伊藤一成: 集合知形成手法を用いた学際的学問分野の知識体系記述の試みーその問題提示と分析ー, 第 1 回データ工学と情報マネジメントに関するフォーラム (DEIM2009) 会議録, C 3-5 (2009).
- [4] "Computing Curricula 2001: Computer Science — Final Report —," The Joint Task Force on Computing Curricula, IEEE Computer Society and Association for Computing Machinery, 236P., December 2001.
- [5] <http://www.loc.gov/catdir/cpsol/lcco/>