

大学情報と情報技術の活用

井田 正明

情報通信基盤の充実とともに社会の諸活動において情報技術の活用が進展している。このことは大学においても同様であり、教育、研究、学務、広報、財務、施設管理などさまざまな分野において急速に情報技術の活用が進みつつある。それらは一般の企業と類似するものであるが大学独自の性質を有するものでもある。昨今、大学では評価に関わる活動が増加しているが評価は根拠資料に基づくものであり、また評価に限らず教育活動の改善の計画策定などにおいても大学情報とその活用は重要である。大学情報の活用は広範囲にわたるが、本稿では組織情報に関するデータベースの活用および教育課程に関する情報（シラバス）の収集と分析について述べる。

キーワード：大学情報、情報収集、情報の分析と提供、データベース、シラバス

1. はじめに

近年、大学の Web ページが充実することにより、一般社会への大学のさまざまな情報の提供が進んでいる。例えば受験生向けの大学紹介情報、大学の組織に関する基礎情報や年度評価などの報告書、シラバスなどの教育情報、また研究に関する研究者ディレクトリや学術情報（論文等）リポジトリなどさまざまな情報が提供されるようになってきた。このような学外への情報発信とともに、学内においても将来計画の策定や内部評価などのためにさまざまな情報の収集と活用が進みつつある。

大学においてはさまざまな組織活動の電子化が進んでいる。研究活動において情報技術は不可欠であるが、教育活動においてもハード面（通信、情報機器）・ソフト面（学習支援ソフトウェア）とも発展し、また組織運営活動においても情報技術の活用が進展している[1]～[3]。大学の諸活動における情報活用は広範囲にわたるが、以下では、組織情報に関するデータベースの活用、教育課程情報（シラバス）の収集とその活用について述べる。

2. 大学情報の収集と課題

近年、大学の教育・研究活動を中心にさまざまな評価活動が行われているが、評価は活動の内容を示す資料・データに基づき評価が行われる。そのため事前に

各大学において評価に必要な資料・データの収集・整備がなされている場合には、評価に際して資料・データに基づく効果的な評価書の作成が可能となりまた作業負担の軽減にもつながる。また、大学の活動を表す大学情報の収集・整備は、評価時に限らず大学の自己改善活動や将来計画の策定など大学の運営にとって重要となる[3][4]。しかしながら、現状は必ずしもそのような状況にはなく、情報の収集・整理は大学にとって負担を伴うものとなっている。

大学情報の収集・整理に関して以下のような課題を挙げることができる。

- データの定義：多様化の進む高等教育の世界では財務情報のようにデータの定義が必ずしも明確なものばかりではない。また全国共通の各種コードの整備も必要である（参考：共通データ CDS、教育プログラム分類 CIP[5]）。
- データ更新と責任問題：定期的な更新とその内容の正確性。
- データ項目の追加変更：頻繁に変更すべきではないが必要に応じて修正するための議論と周知方法。
- 文書情報：各種議事録、規定、シラバスなど数値情報以外の文書データの収集と管理。
- データのレベル：教員や学生など個人データから必ずしも木構造となっていない組織構造での集計。
- データの収集の負担。
- 大学情報を管理し所掌する組織の必要性と人材の育成：これまでの教職員の範疇にはなかった大学の情報を使いこなせる人材の育成体制（ア

いだ まさあき
大学評価・学位授与機構 評価研究部
〒187-8587 小平市学園町西 1-29-1

- ・ メリカの専門職団体[6].
- ・ データベースの分散化[4].

3. 組織情報に関するデータベースの活用

ここでは大学に関する国レベルのデータベース[7][8]について考える。日本においては、学校基本調査などの文部科学省をはじめとする政府関連の統計調査があるが、公表されている情報は例えば都道府県別であるなど集計情報であり各組織単位の詳細なものではない。科学技術振興機構の研究開発総合ディレクトリでは、研究者情報や研究課題情報などを収集・提供している[9]。その他に、大学入試センター、国立情報学研究所、メディア教育開発センター、またいくつかの大学において全国的な高等教育に関する情報が提供されている。国内の多くの大学では、研究業績に関する教員情報データベースや学術情報レポジトリの構築が進みつつある。

海外における国レベルのデータベース (National Dataset) として、アメリカ合衆国の事例を述べる。アメリカにおいて高等教育に関するデータベースは多数存在し、またその提供情報は詳細である。NCES (National Center for Education Statistics) はさまざまなデータを提供しているが、その中でも IPEDS (Integrated Postsecondary Education Data System [10]) は全米の高等教育機関を対象とした包括的なデータベースシステムであり、高等教育機関ごとに学生の入学、卒業や財務、職員給与などさまざまなデータが収集・集積されている。IPEDS はインターネットを介してデータの取得や分析ツールの利用が可能のため、国、州、各高等教育機関レベルでの高等教育の状況を把握するために活用することができる。NCES 以外にも NSF (National Science Foundation) も高等教育に関する情報を Web で提供している。分析システムとしては Webcaspar を挙げることができる [11].

以下では IPEDS についての紹介を行う。各大学から IPEDS へのデータ入力項目はつぎである：12-Month Enrollment (E12), Completions (C), Fall Enrollment (EF), Finance (F), Graduation Rate (GRS), Human Resources (HR), Institutional Characteristics (IC), Student Financial Aid (SFA)

また、州によっては IPEDS データへの追加・補完として別途データが収集されている。

IPEDS が Web を介して提供するデータツールはつ

ぎである。

- ・ Peer Analysis System (PAS)：基準となる大学と比較対象となる大学集合との比較結果を提供。
- ・ Executive Peer Tool (ExPT)：PAS の簡略版。
- ・ Dataset Cutting Tool (DCT)：IPEDS データをカスタマイズして提供。
- ・ College Navigator：全米約 7,000 大学についてのディレクトリ。受験生などへ大学の詳細な情報を提供 (複数大学の比較も可)。
- ・ Data Analysis System (DAS)：さまざまな分析を行う解析テーブル (パラメータ集合) を生成。

PAS：注目する大学と他の大学との比較分析は PAS を利用することになる。比較対象とする大学を個別に選択することやシステムが提供する分類 (カーネギー分類など) によって選択することが可能である。選択後に図 1 に示すように比較するデータ項目を決定する (例えば、授与学位数や教育・研究に関する費用等)。比較結果のテーブルは Web 画面で容易に確認することができる (結果を CSV 等のファイルにダウンロードすることも可能)。この PAS システムではデータ項目を細かく選び比較検討を行えるが、これを簡略化した分析システムが ExPT である。事前に用意された基本的な分析指標について他大学群との比較分析を行うことができ、さらに結果はグラフ化され容易に結果を確認することができる (図 2)。また、IPEDS に収められているデータから必要項目を抽出する DCT も用意されている。

DAS：IPEDS のデータから分析テーブルの生成と共分散分析を行うため、変数に対して条件を詳細に与え

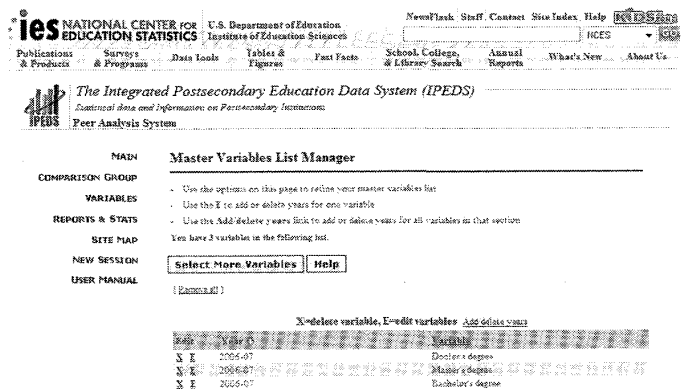
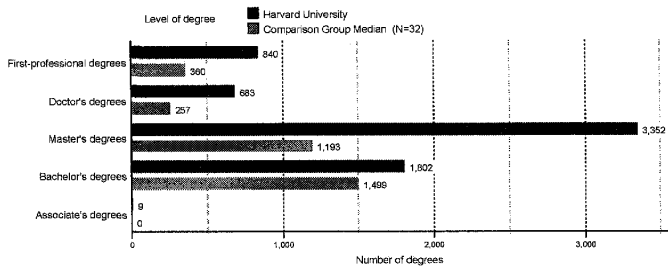


図 1 分析項目の選択
(データ出典：NCES；<http://nces.ed.gov/ipeds/pas/>に具体的データを入力し項目を選択)

Number of degrees awarded, by level: Academic year 2006-07



U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2007, Completions component.

図2 ExPTのグラフ出力
(データ出典: NCES; <http://nces.ed.gov/ipeds/spas/ExPT/>に具体的データを入力した出力情報: 授与学位数の比較)

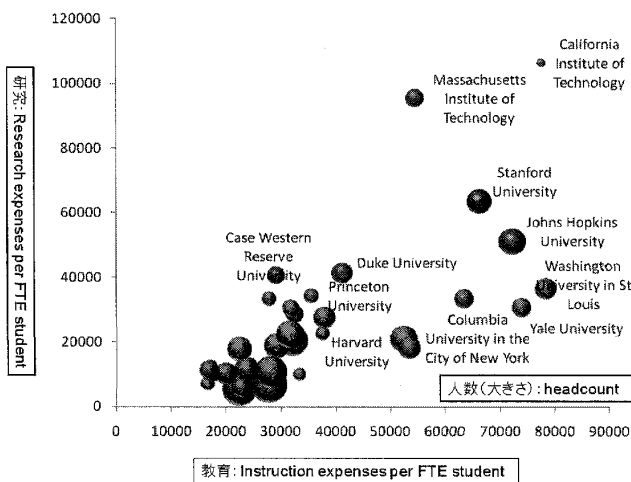


図3 ダウンロードしたデータの分析例
(データ出典: NCES; <http://nces.ed.gov/ipeds/spas/ExPT/>より取得した教育・研究費用のグラフ化)

ることができる。このプログラミング指示ファイルは TPF (Table Parameter File) とよばれ複雑なパラメータ集合を作成し保存することができるため詳細な分析が可能である。

以上のように IPEDS ではデータ分析ツールによって Web ベースで容易に分析テーブルが作成できる。また分析結果のテーブルは CSV や SPSS などの統計ソフトのファイル形式でもダウンロードができ利便性のよいものとなっている (図3)。

これら分析ツールは簡便な利用から詳細で高度な利用まで可能であり、複雑な利用を行うためには技術が必要となる。また IPEDS の分析システムに限らず大学のデータの分析には専門知識が必要とされる。このような情報の収集や分析を担当する専門職として多くの米国の大学では、ネットワークや情報システムの運

用など情報の技術面を担当する部署とは別に、その大学自身についての情報の収集や分析といったことを専門に担当する部門を設けている。これらは、Office of Institutional Research (IR 部門) などの名称でよばれている。IR 部門の業務は、その大学の状況に関する情報を収集し分析を行うことである。またこれら専門職の教育プログラムも存在する [12]~[14]。

4. 教育課程情報の収集と活用

4.1 シラバス情報の収集

教育活動の電子化については情報通信基盤の整備やソフトウェア開発で過去数多くの事例がある [1][2]。教育活動に関してもさまざまな情報が大学において収集されている。特に工学系では JABEE との関連で教育情報が詳細に蓄えられる事情がある。今後さらに教育情報の収集の必要性を考える必要がある (海外との関係や分野別の質保証等)。

ここでは教育課程情報としてシラバス情報の収集について考える。数多くの大学の Web ページよりシラバスの構造特性を利用して自動抽出を行う研究としては文献 [15] や [16] を挙げることができる。また、メディア教育開発センターでは、学習コンテンツの総合的なゲートウェイとして NIME-Glad を開発している [17]。これは、多くの e-Learning コース、学術講演・公開講座、シラバスなどの学習コンテンツを体系的に提供するものである。シラバスについては多数の大学のシラバスを集約してデータベースを構築する。そのためインターネット上に分散しているそれぞれのコンテンツに LOM (Learning Object Metadata) を付与しデータベースに登録している (タイトル, 提供場所, 概要, キーワード, 教育分野, 分類等)。学習者は、LOM のデータベースを検索することによって学習コースを見つけることができる。

4.2 シラバスの分析と活用

多数のシラバスや履修科目表等の教育課程情報を収集しそれらの内容を分析する研究開発も進みつつある。ここではシラバスより内容情報を抽出・整理した上でデータベース化し、テキスト分析することにより、各シラバスに含まれる専門用語に基づいて教育課程を比較検討する研究について述べる [18][19]。これにより各大学・学部の教育課程の分析、比較、設計などが進展するものと期待される。現在、日本の大学においては、学士課程の専門教育では数十科目に及ぶ科目が開講される。単一の学科については、そのシラバスを大

学教員などが直接読むことで教育課程の編成状況を理解できるが、類似の教育目標を持つ多数の教育課程について、シラバスをすべて読み相互比較することで特徴を把握することは困難である。

シラバスデータのクラスタリングに基づく教育課程の分析システムにおける分析手順はつぎのようになる[18]：(1)分析対象とする教育課程とシラバスを選択，(2)専門用語抽出を適用するデータ項目を選択，(3)含まれる専門用語に基づき各シラバス（科目）の内容を定量化，(4)シラバス間の類似度を計算，(5)類似度に基づきクラスタリングを実施，(6)クラスタ数を選択，(7)クラスタ要約の分析，(8)コレスポネンス分析と結果の2次元または3次元表示による近接性（クラスタ-クラスタ間，学科-学科間）の把握，(9)分析目的に応じて(1)～(6)のステップに戻り再分析を実行。

これにより教育課程の分析が行われる。例えば，Web上にシラバスが公開されており学科名が「システム工学」を含んだ（多様な授業科目を有する）13学科。それらの学科が提供している教育課程のシラバス集合を分析対象としたものを考える[19]。システム工学は領域横断的な性格を有し具体的な適用領域と融合した教育課程が数多く見られる。システムを名称に含む学科は，例えば，機械-，化学-，情報-，経営情報-，機械知能-，応用生命-，福祉-，など多岐にわたる。

分析システムの各種設定は次のようにした。専門用語抽出の対象項目：科目名，授業概要，授業目的，授業計画，教科書。専門用語辞書：EDR。用語重み付け関数：TF-IDF 値（Term Frequency-Inverse Document Frequency）。シラバス間類似度：コサイン類似度，クラスタリング：群平均距離を用いた階層併合的クラスタリング。クラスタ数を7としてクラスタリング。

学科（大学）を各行，クラスタを各列にしたクロス表を表1に示す（表中の数値はシラバス数を表している）。学科ごとにそれぞれのクラスタに属するシラバス数を見ることで，各学科がどのクラスタに関連が強いかを検討することができる。

この表をコレスポネンス分析し2次元表示したものが図4である。

図4の中で，C0は機械工学，化学工学，システム工学，制御工学などの分野の用語を多く含むクラスタである。学科U07（機械システム工学科），U35（化学システム工学科），U73（機械システム工学科），

表1 クラスタリング結果のクロス表[19]

	C ₀	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	計
U07	10	1	3	0	0	0	0	14
U09	7	10	6	2	10	2	0	37
U13	17	14	13	29	3	2	0	78
U15	10	5	8	8	6	0	0	37
U19	5	8	6	4	19	2	0	44
U32	4	1	11	1	1	0	0	18
U35	58	0	2	5	3	0	0	68
U37	2	10	24	16	5	3	0	60
U50	46	2	8	2	23	1	1	83
U61	25	7	16	6	7	1	0	62
U63	7	9	18	9	8	2	0	53
U73	31	1	4	5	0	0	0	41
U78	32	1	3	1	4	0	0	41
計	254	69	122	88	89	13	1	636

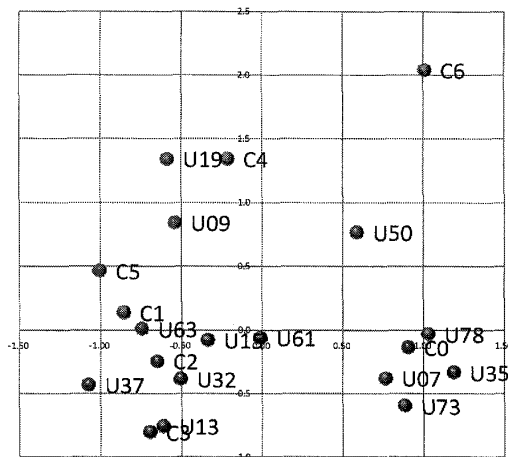


図4 コレスポネンス分析

U78（機械システム工学科）がこのクラスタに近接し，それら学科の提供する教育課程はそれらの用語を多く含む科目で構成されている。また，C1はプログラミング，C2は通信，C3は電気回路，C5は画像やパターン処理に関する用語を多く含むクラスタである。学科U13（応用システム工学科），U15（システム工学科），U32（システム工学科），U37（情報システム工学科），U61（システム工学科），U63（知能情報システム工学科）の教育課程はこれらのクラスタにより近い。C4は数学的な科目を多く含むクラスタである。学科U09（情報システム工学科），U19（情報システム工学科）の教育課程はこの方面にも関係が深いといえる。C6は実験に関する用語が含まれており，学科U50（機械システム工学科）が有する実験科目のシラバスとの関係を見ることができる。以上のように，学科名とクラスタの関係は大局的にみて常識的な傾向を見てとることができ，また，一部に特徴的な学科を見

てとることができる。

一般に、個々のクラスタに対して意味づけができないこともある。これはシラバスのような文書データを対象としたクラスタリングでは、分析を適用する文書の質の相違（記述量のばらつきをはじめ、シラバス執筆者ごとの使用する用語や文章表現の違いなど）が結果へ影響を及ぼすためと考えられる。しかしながらシラバスの多量な文書集合を用い分析を行っていることから対象集合の全体的な傾向は捉えることができ、これにより各学科の特徴把握は可能といえる。

このような情報技術を活用した詳細な教育課程の分析が進むことにより、各大学における教育課程の設計や他大学との比較検討が進むものと予想され、今後、さらにシラバスを利用した計算機支援法の研究開発は重要性を増すものと考えられる。その際に利用される電子データはさまざまな大学から Web 等より収集するため、データの形式が非統一であり、データベースの構築に多大な負担がかかるため、シラバスの基本的構造の標準化が必要とされる。

その他にシラバスを分析対象とした研究開発の例として、東京大学工学部では工学教育推進機構を設置しており、その中で工学知の構造化教育を進めている[20]。そこでは工学教育の講義・演習課目を工学知の粒度と考え、シラバスを基にして工学系カリキュラムの構造化と可視化を行っている[21]。これにより知識（カリキュラム）の全体像を可視化し学生が進むべき道を自ら検討し選択が可能となる。また知識の重複や不十分な箇所を発見することができる。さらには全体像より知識発見につなげていくことが考えられ今後さらなるシラバスの活用が考えられる。またシラバスにオントロジーを用いた知識の記述による講義構造の構造化を行うことによってより高度な検索を行える講義データベースの研究もある[22]。

5. 大学情報の分析と提供

大学情報の分析と提供は様々な状況で行われる。

- ・ 分析と提供を行う主体としては、大学（IR 部門）、国（文科省など）、高等教育研究者、企業（出版・教育等）が考えられる。
- ・ 分析された情報の提供先としては、大学受験者（父母、高校）、企業（就職先）、企業（共同研究）、監督官庁、評価機関、国民が考えられる。

前述した大学内における IR 部門の役割の大まかな 4 分類が考えられている[13]。IR 部門を分析と情報

提供の主体と考えると、4つの分類は、情報の提供方向を大学内方向と外方向に分類し、また分析内容を基準にした分類とみなすことで次のように解釈ができる（{内, 外}×{事務的, 研究的}）。

- ・ 自大学の説明のため（情報の権威者としての IR）：学内における必要な情報を蓄えており、必要時にはデータはいつでも提供できる。学内データベースとしての役割。情報管理と技術サポート。
- ・ 選択肢の分析のため（政策分析者としての IR）：計画立案や意思決定をするために学内のデータを分析する役割。年度計画、将来計画策定を支援。
- ・ 最良の事例を提供するため（報道対策助言者としての IR）：広報活動、学術情報リポジトリ、教員データベース、共同研究・地域連携への支援。
- ・ 有効性の公正な証拠の提供のため（研究者としての IR）：高等教育に関係する高度な分析。

これ以外には例えば、主体として出版企業が大学ランキングを受験者（一般）対象で出版提供する場合が考えられる。また高等教育研究者が国内の高等教育の情報を分析し全国的傾向を一般国民や監督官庁に提示する場合も考えられる。

前述の IPEDS はその役割として、各大学の IR 部門が自大学と他の比較対象大学を比べることに利用したり、高等教育研究者が分析に用いたり、また受験生が複数の大学の基礎情報を比較して志望校を絞り込む等に利用される。

大学情報の分析として多くの人に関心を持たれるのは大学ランキングであるが研究に於ける各種指標比較も興味を持たれている。データを分析する際には、指標の妥当性、データの定義（の不正確さ）、組織比較（組織規模の違い、分野の違い）などに注意を払う必要がある。意味ある分析を行うためには、分析元データが公開され、分析指標の設定が適切であり、またその分析結果の意味を適切に解釈できる専門家が必要と考えられる。

6. おわりに

昨今、国立大学法人化、グローバル化による国内外の競争的環境への変化、産業界の大学に対する期待の変化など大学をとりまく環境は急速に変化してきた。このような環境の中で大学の諸活動を支援するシステ

ムのための高度な情報技術に関係する研究・開発は、今後ますます重要になるものと考えられる。この新たな研究・開発分野はテキストマイニング、統計技術、高等教育に関する知識など多様な知識が求められる研究分野であり、今後の発展が期待できるものと思われる。

参考文献

- [1] 経済産業省商務情報製作局情報処理振興課 (編), eラーニング白書 2007/2008 年版 (2007).
- [2] 坂本昂: 人類の知を有効活用する e-Learning, 日本知能情報ファジィ学会誌, Vol. 20, No. 6, pp. 65-77 (2008).
- [3] 井田正明, 喜多一: 高等教育機関の諸活動と情報技術, 日本知能情報ファジィ学会誌, Vol. 17, No. 5, pp. 3-10 (2005).
- [4] 井田正明: 評価のための資料・データの収集と活用, 大学評価・学位授与機構編著: 大学評価文化の展開—評価の戦略的活用を目指して, ぎょうせい, pp. 13-21 (2008).
- [5] Common Data Set Initiative (CDS), <http://www.commondataset.org/>, および Classification of Instructional Programs (CIP 2000), <http://nces.ed.gov/pubs2002/cip2000/>
- [6] Association for Institutional Research (AIR) <http://www.airweb.org/>
- [7] 個別大学情報の内容・形態に関する国際比較, 東京大学 大学総合教育センターものぐらふ, No. 2 (2003).
- [8] 高等教育データベースを用いた分析の試み, 東京大学 大学教育総合センターものぐらふ, No. 5 (2005).
- [9] ReaD: <http://read.jst.go.jp/>
- [10] IPEDS: <http://nces.ed.gov/ipeds/>
- [11] WebCASPAR: <http://webcaspar.nsf.gov/>
- [12] Knight, W. E., Ed.: The Primer for Institutional Research, Resources in Institutional Research, No. 14, Tallahassee, FL: Association for Institutional Research (2003).
- [13] Volkwein, J. F., Ed.: What Is Institutional Research All About? A Critical and Comprehensive Assessment of the Profession, New Directions for Institutional Research, No. 104, San Francisco: Jossey-Bass Inc., Publishers (1999).
- [14] Institutional Research Certificate Program: <http://www.ed.psu.edu/ircert/ir.html>
- [15] 山田信太郎, 松永吉広, 伊東栄典, 廣川左千男: Web シラバス情報収集エージェントの試作, 電子情報通信学会論文誌, J 86-D-I, 8, 566-574 (2003).
- [16] 地蔵真作, 篠原正典: 高等教育授業シラバスの Web からの自動抽出—自然言語処理を利用した自動抽出法とその結果—, NIME 研究報告, 25 号 (2007).
- [17] 清水康敬: 学習対象メタデータ (LOM) 付与による教育用コンテンツの共有と流通, 情報処理, Vol. 46, No. 6, 677-683 (2005); NIMEglad: <http://nime-glad.nime.ac.jp/>
- [18] 野澤孝之, 井田正明, 芳鐘冬樹, 宮崎和光, 喜多一: シラバスの文書クラスタリングに基づくカリキュラム分析システムの構築, 情報処理学会論文誌, Vol. 46, No. 1, pp. 289-300 (2005).
- [19] 井田正明, 野澤孝之, 芳鐘冬樹, 宮崎和光, 喜多一: シラバスデータベースシステムの構築と専門教育課程の比較分析への応用, 大学評価・学位研究, No. 2, pp. 85-97 (2005).
- [20] 工学知の構造化教育: <http://ciee.t.u-tokyo.ac.jp/ciee/index-ciee.html>
- [21] 大場善次郎: 「工学知の構造化と可視化」の試み—工学教育に向けて—, 大学評価・学位研究, Vol. 1, 97-109 (2005).
- [22] 関谷貴之, 山口和紀: 講義データベース—オントロジーを用いた要求知識と習得知識の記述による講義情報の構造化—, 日本知能情報ファジィ学会誌, Vol. 17, No. 5, pp. 22-32 (2005).