共通テストデータを要約した簡易統計情報検索システムの開発

石岡 恒憲[*]
鈴木 規夫[*]

要 約　共通テスト（共通第一次学力試験および大学入試センター試験）の志願者数、個別試験の志願者数、およびその合否状況について、昭和54年度から現在までの時系列推移をユーザが指定するさまざまな表示分類（学部系統別、性別、現浪別、高校課程別、高校学科別、受験先別、およびこれらの組み合わせ）にしたがって検索/表示できるシステムを試作した。単に数値だけでなく、Excelによって作成された明瞭な図が提示される。これらの情報はWebにより提供され、ユーザは特別な設定を必要とせずに、簡単に検索をおこなうことができる。

キーワード：Apache, cgi, Excel, HTML, Web, Windows, WWW, 共通テスト志願者数、個別試験志願者数、合否状況

---

[*] 大学入試センター 研究開発部 試験作成支援研究部門
〒153 東京都目黒区駒場 2-19-23　e-mail: tunenori@rd.dnc.ac.jp　Phone: 03-3468-3311

1　はじめに

共通テスト（共通第1次試験および大学入試センター試験）が実施されて20数年が経過した。この間、志願者に関する膨大な情報が蓄積されているにも関わらずその情報を電子媒体で容易に検索し入手する手段が現在、存在しない。

現在、紙媒体で公開されている詳細で一般に入手可能な冊子は、大学入試センターが毎年発行する大学入試フォーラムである。ここには、当該年度の各選抜（一般選抜、専門高校・総合学科卒業生選抜、アドミッションオフィス入試選抜、特別選抜）毎に、国公私立大別に募集人員、入学志願者数、志願倍率、受験者数、合格者数、入学者数、過欠員数が掲載されている。また、大学入学状況の推移としては、時系列の前年度高校卒業者数、現浪別の入学志願者数、入学者数、入学率、（入学者の）同一年令比が掲載されている。

同じく大学入試センターより刊行される「大学入試センター試験— 実施結果と試験問題に関する意見・評価 —」と題される冊子には、出願資格別の志願者数や、高校卒業見込者・卒業者の学科別内訳、および男女別内訳があり、さらに都道府県別の志願者数内訳が掲載されている。

このように志願者数や合格者数等についての最も基礎的な情報は幾つかの冊子に掲載されているが、それ以外の情報についてはどこに、どのような情報が存在しているかを容易に知ることができない。また志願者に関する属性は大別して、本人の属性（性別、現浪別）、志願先の属性（学部系統別、受験先）、出身高校の属性（課程別、学科別）の3つに分類されると考えられるが、それぞれの分類、およびそれらの組み合わせによって、志願者情報を観察、あるいは解析することは、データ解析における層別の観点からも非常に重要なことである。たとえば、岩田・山田(1993)は学部系統別にみた国公立大学志願者層の推移について調査している。また鈴木(1984)は性別と出願学部系統と受験年数を主因子とし、受験行動に与える影響について調査している。

しかしながらこのような研究は、研究者自身が、必要な度に大量のデータ処理を行って解析結果を得たものであり、入試研究をより活発化させる意味でも、できるだけ情報の集約化を計るべきであろうと考える。また研究開発部だけでなく、事業部や管理部等で共通に活用できる検索システムの構築が望まれる。

そこで著者らは、過去20数年間の情報を要約した統計情報ファイルを作成し、

- それらの情報が即座に検索でき、
- それを図として表示し、
- 利用者にわかりやすく提供できるシステムを開発する

ことにした。

著者の一人は以前に、たとえば鈴木(1984)にあるように過去のデータについての志願者統計情報を作成しており、どのような情報の蓄積が必要であるかのある程度の経験をもっている。この経験に基づき、著者らがまず検索システムを試作し、これをベースにして各課とも相談しながら、ニーズに沿った統計情報の提供をめざしたいと考えている。

また将来に渡っての基礎的情報の保存という観点からも、志願者に関するデータベース化は必須の技術課題である。本システムはそのための手始めとしたい。

2　システム概要

2.1　本システムの特徴

我々が提案するのは、志願者マスタ（データ）から中間的な統計情報を要約し、その情報をベースにして、共通テスト志願者数等の年次変化の様子を属性（性別、高校課程、高校学科、個別試験の志願学部系統等）の違いを加味しながら、検索することのできるシステムである。検索上での多少の制限は残ることになるが、格納されるデータ量は少量で済み、短時間で容易に結果を得ることができる。

他の統計情報システム、たとえば総務省統計局統計センター `http://www.stat.go.jp/` が公開する国勢調査抽出集計統計表や人口推計な

どの統計は、その種類は多いものの全て Excel のファイルとして提供されているに過ぎない。これに比べ我々のシステムは、

- 図が即座に表示され、全体としての傾向の把握が容易であること
- 属性に応じたクロス集計ができること

の特徴がある。

ユーザ（クライアント）からは Web によって検索可能なシステムとする（図 1）。ユーザは特別な設定なしに、自分のマシンから Netscape や Internet Explore 等の好みのブラウザで利用できる。JavaScript は必要としない。現在は、センター内（dnc.ac.jp ドメイン）からのみ、検索可能となるよう設定をしている。また Web サーバには、Windows システムの標準サーバである IIS (Microsoft Internet Information Server) を用いずに、セキュリティ上、頑健な Apache を用いているのも特徴の一つである（昨年 2001 年末に前世界中に猛威をふるった Nimda ウィルスは、IIS のセキュリティ上の脆弱性を狙ったものである）。

サーバ側では、Perl によって書かれた cgi プログラムによって、HTML 形式で記述された図（共通テスト志願者数等の数値データを時系列データとして図に示したもの）およびテキスト（図作成のもととなる数値データ）を検索結果として返す。本システムでは検索結果をその都度、動的に作成するのではなく、全ての検索結果の図表を予め用意しておき、ユーザが指定する検索条件に応じて該当する結果の図表を提示する。これにより、検索におけるユーザの待ち時間を大幅に減じることができる。

作成すべき図表は全部でそれぞれ数千個であるが、Excel のマクロ (VBA) の機能を利用して作成したプログラムを実行することでこれらを自動的に生成できるようにした。

2.2 検索/表示項目

本システムで検索/表示できる項目は、以下の 3 つ、あるいはその組み合わせである。

1. センター試験志願者数
2. 個別試験志願者数（出願）
3. 合否状況

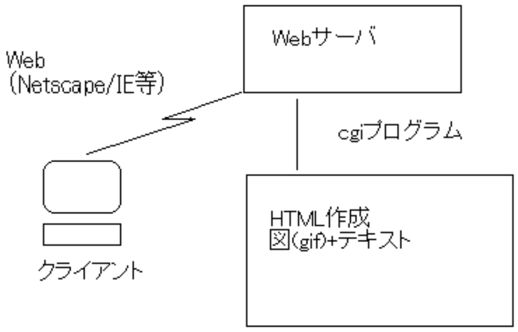


図 1: システム・フロー

1. については「出願者全体」、「受験者/未受験者別」、「受験者のみ」を選択できる。

2. は「出願者のみ」を対象とする。

3. については「合格者/辞退者/不合格者別」、「合格者のみ」、「辞退者のみ」を選択できる。

検索/表示する際に、さらに以下の 6 つの属性分類を、最大 2 つまでの組み合わせで指定できる。

1. 学部系統別
2. 性別（男/女）
3. 現浪別（現役/浪人）
4. 高校課程別
5. 高校学科別
6. 受験先（国私/国立/私立）

「1. 学部系統別」については「人社/理工/農水/医歯薬/保健/教員/その他/II 部」の全学部系統ごとに分類して表示する方法と、「人社」、「理工」、「農水」、「医歯薬」、「保健」、「教員」、「その他」、「II 部」といったように特定の学部系統のみを対象とすることができる。これは学部系統ごとの違いを把握したいという要求の他に、個別の学部系統に着目した動向を把握したいという要求に応えるための配慮である。学部系統の分類については付録 C を参照されたい。

「4. 高校課程別」については「全日/定時/通信/高外」の全課程について分類/表示する方法と、「全日」のみを対象とするする方法、「定時/通信/高外」の全日以外を対象とする方法、「高外:外国/高専/在外/指定/大験」の高校外のみを対象とする方法が選択できる。

高校課程別では、「全日」課程がセンター試験の受験において主流であるので、これに着目した選び方と、これを除く選び方、さらに例外であるところの高校外のみに着目した選び方ができるよう工夫した。全ての課程を一つにまとめてしまうと、特に高校外においては「外国/高専/在外/指定/大験」の区別が図では判読できなくなってしまう。

「5. 高校学科別」については「普通/理数/農業/工業/商業/総合/高外」の全学科について分類/表示する方法と、「普通」のみを対象とするする方法、「理数/農業/工業/商業/総合/高外」の普通を除く学科を対象とする方法が選択できる。

これはセンター試験の受験において主流であるところの「普通」に着目した選び方と、これを除く選び方ができるようにとの配慮である。

3　検索例

以下は本検索システムにおいて、典型的な、あるいは特徴の顕著な検索事例の結果である。システムの操作方法の詳細については、付録 A を参照されたい。

3.1　「学部系統別」にみた「センター試験の出願者数」の推移

昭和 54 年度当初、5 教科 7 科目であった共通 1 次試験は、昭和 62 年度に 5 教科 5 科目となった。平成 2 年にはセンター試験が開始し、私大が参加するようになった。受験科目においてもいわゆるアラカルト方式が採用された。

現在、センター試験の出願者は約 60 万人であるが、国公立大学への志願者に限れば、その数は昭和 54 年度の約 34 万人から変わらずほぼ一定である。そこでその内訳を把握するために「学部系統別」、すなわち「人社系/理工系/農水系/医歯薬系/保健系/教員系/その他/II 部」の 8 つの学部系統ごとに「センター試験の出願者数」を示したのが図 2である。平成 12 年度において人数の多い方から順に理工系、人社系、教員系、医歯薬系、その他、農水系、II 部、保健系である。

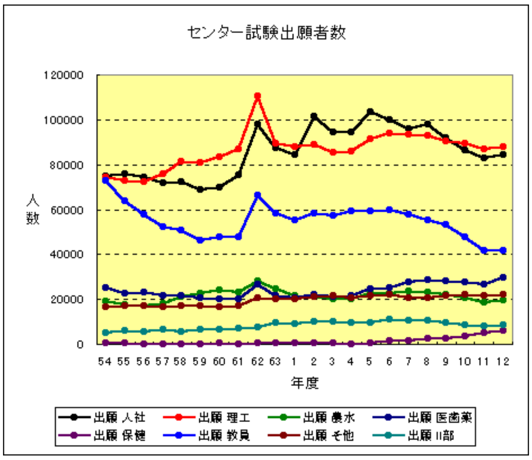


図 2: 「学部系統別」にみた「センター試験の全出願者数」

これより

1. 昭和 62 年度の志願者が突出して多いこと
2. 平成元年から 8 年あたりまで、いわゆるバブル経済のときには人社系の志願者が多かったが、近年は減少傾向にあること
3. 理工系は昭和 63 年以降、志願者数が一定で、最新の平成 12 年度においては人社系を抜いて志願者最大の学部系統となったこと
4. 教員養成系は、昭和 54 年度においては、人社系、理工系と並ぶ出願数であったが、近年の就職難を反映してか、志願者が減少傾向にあること
5. 近年は、医歯薬系、および保健系の志願者が増加していること

などがみてとれる。

昭和 62 年度は、この年より 5 教科 5 科目となり、また複数回の受験が可能となったのであるが、この年のみが顕著な振る舞いをみせるの

は、この年のみがセンター試験の結果を待たずに個別大学への出願を行なういわゆる事前出願が実施された年であることに起因すると考えられる。（事前出願は翌 63 年には廃止された。）

3.2 「高校外課程」における「センター試験の出願者数」の推移

高校の課程を経ない、いわゆる「高校外課程」における「センター試験の出願者数」の推移について示したのが図 3である。このデータは教育社会学の立場から興味深いものと思われる。平成 12 年度において人数の多い方から順に大検（大学入試資格検定）、高専（高専 3 年修了）、在外（在外教育施設）、外国（外国の学校等）、指定（文部大臣の指定した者）を示す。

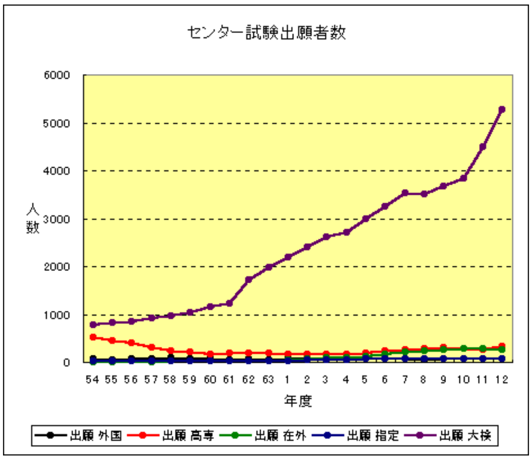


図 3: 「高校外課程」における「センター試験の全出願者数」

これより

1. 大検資格による出願者の伸びが顕著で、昭和 54 年から平成 12 年度までで約 6 倍に達し 5,000 名を越えていること
2. 高専修了資格による出願者は、昭和 54 年以降むしろ減少しており、近年でも増加していないこと
3. 数的にはそう目立たないが、昭和 54 年ではわずか 2 名であった在外教育施設の出願者が現在では 200 名を越えていること

などがわかる。

3.3 「医歯薬系」における「合否状況」の推移

「医歯薬系」における「合否状況」の推移を示したのが図 4である。平成 12 年度において人数の多い方から順に不合格、合格、辞退を示す。

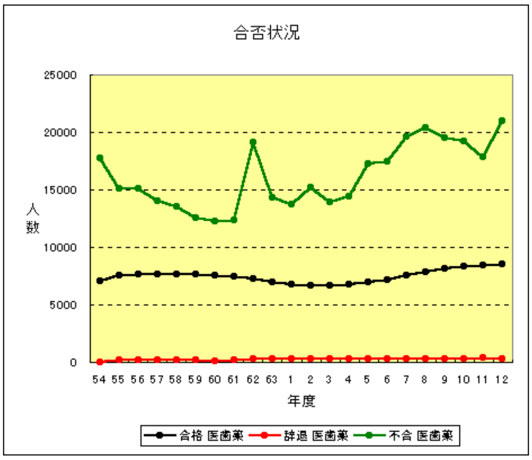


図 4: 「医歯薬系」における「合否状況」

これより以下のことがわかる。

1. 昭和 62 年度の不合格者が突出して多いのは、その年の入試制度が特異であったことによる影響が大きいと考えられようが、平成 2 年度以降、不合格者数が増加しており、合格者数がその間、一定である（すなわち入学定員が増加していない）ことを考えれば、近年、医歯薬系が受験生にとって人気であるとみなすことができる。
2. 平成 2 年以降、辞退者がほぼ一定で 300 名程度いる。

本節で示した事例以外にも、着目したい事項に応じて、多くの検索結果を容易に得ることができる。

4 システム運用

4.1 データ・メンテナンス

毎年作成される志願者マスタから、SAS(PROC FREQ) によって、検索条件となる全ての属性を含んだ中間的な集計データファイルを作成す

る。次にこの中間データファイルについてExcelのVBAプログラムを動作させ、検索の際に指定するパラメータに応じた、図、および図作成のもととなる数値データを作成する。これらはExcelのマクロ機能を利用することで、自動的に生成できるようにしてある。

図はgif形式で作成されるので、検索結果として返すHTMLファイル中には、これらgifファイル名を埋め込む形で記述している。したがってこの新しく作成したgifファイルをWebサーバの適当なディレクトリに転送するだけで済む。

一方、図作成のもととなる数値データは、HTMLのtable形式に変換する必要がある。数値データをそのままテキスト形式で格納したのでは、桁があわずに見苦しいものとなるからである。著者らはこのためにデータ成形用の簡単なスクリプトを作成した。このスクリプトでは、Excelの数値データを単にHTML形式で保存するだけでは実現できない項目ラベルを追加し太字にする、また数字データを右揃えにするなど、見やすくするための幾つかの工夫を施している。さらにスクリプト起動時に、そのときの日にちと、現在処理しているデータが平成何年度までのデータであるのかを自動的に判別し、これを「データ更新情報」に表示されるようにした（付録、図9）。これによりユーザは、自分の行なう検索の対象期間を検索開始前に分かるようになるだけでなく、万一、データに間違いが発見されてデータ修正が必要となった場合に、データ修正がシステムに反映されたか否かが（検索前に）わかるようになる。

この成形された数値データは、その後Webサーバの適当なディレクトリに転送することになる。

これら一連の作業までもが全て自動というわけではない。しかしながら必要な作業単位はほぼ自動化されており、また作業が年に1回であることを考えれば、データ更新に伴うメンテナンスにかかわる労力は軽微なものと考える。また操作マニュアルを作成した。

4.2 マシン環境

現在のマシン環境、ならびにシステムに用いている言語、ライブラリ、応用ソフトウェアは、以下の通りである。

Webサーバ: Sony VAIO PCG-R505 (Windows 2000 Professional) + Apache 1.3.14
http://www.apache.org/
Apacheは現在、インターネット上で圧倒的に多く使われているフリーのWebサーバ・ソフトである。

cgi: Perl v5.6.0
http://www.ActiveState.com/ActivePerl/
ブラウザからの入力を受けてWebサーバ上で動作するプログラムはPerlによって作成した。

cgi-lib.pl: Version 2.18
http://cgi-lib/berkeley.edu/
フォーム入力を扱う上で役に立つPerl言語用の入力ユーティリティを用いた。

Microsoft Excel: 検索結果の図を作成するのに使用した。図はgif形式で格納される。

5 今後の方針

共通テスト志願者数、ならびに個別試験志願者数、およびその合否状況について、すなわち出願から受験、および結果についての最も基本的な統計諸量について、昭和54年度から現在までの推移をユーザが指定するさまざまな表示分類にしたがって観察できるシステムが試作できたのではないかと考えている。今後は以下について検討を加えたい。

1. 検索情報の検討

   現在、事業部、管理部にもURLを公開の上、運用中である。

   今後は、どのような属性が検索対象として必要かを関連部署との間で検討しながら仕様をつめてゆきたい。

2. 検索フローについての検討

検索から表示、およびおよびその操作の繰り返し等、一連の検索フローについて不備/使いにくさがないか、検討を加える。

謝辞

論文を改訂するにあたり有益なコメントをいただきました 2 名の匿名の査読者に厚くお礼申し上げます。

参考文献

大学入試センター (2001): 2001 大学入試フォーラム, No.24.

大学入試センター (1999): 平成 11 年度 大学入試センター試験— 実施結果と試験問題に関する意見・評価 —

Gundavaram, Shishir（監訳）田辺 茂也 (1996): CGI プログラミング, オライリー・ジャパン.

岩田弘三, 山田文康 (1993): “学部系統別にみた国公立大学志願者層の推移に関する研究,” 大学入試センター研究紀要, No.22, pp.1–52.

ローラ・リメイ, アーマン・ダニッシュ（訳）武舎 広幸, 久野 禎子, 久野 靖 (1998): 続・HTML 入門第 2 版 —新機能,CGI,Web の進化—, ピアソン・エデュケーション.

鈴木規夫 (1984): “共通第 1 次学力試験からみた大学学部の出願に関する決定過程とその成否,” 大学入試センター研究紀要, No.7, pp.1–25.

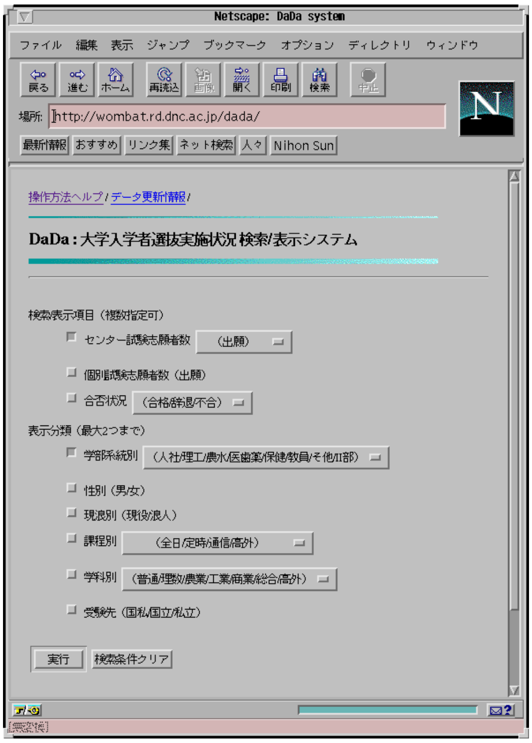
付録

A システム操作方法

ユーザは Netscape や Internet Explore など好みのブラウザを立ち上げ、

`http://dada.rd.dnc.ac.jp/dada/`

と入力する。すると図 5 のような画面が現われる。



場所: http://wombat.rd.dnc.ac.jp/dada/

DaDa : 大学入学者選抜実施状況 検索/表示システム

検索表示項目（複数指定可）
☑ センター試験志願者数 （出願）
☐ 個別試験志願者数（出願）
☐ 合否状況 （合格/辞退/不合）

表示分類（最大2つまで）
☑ 学部系統別 （人社/理工/農水/医歯薬/保健/教員/その他/II部）
☐ 性別（男/女）
☐ 現浪別（現役/浪人）
☐ 課程別 （全日/定時/通信/高外）
☐ 学科別 （普通/理数/農業/工業/商業/総合/高外）
☐ 受験先（国私/国立/私立）

図 5: 検索開始画面

操作 1: 検索/表示項目を指定する。

複数を指定した場合は、指定した全ての表示項目について表示する。特に指定しなければ（初期状態では）「センター試験志願者数」を表示する。図 5は検索開始時における初期状態での画面である。

操作 2: 表示分類を選択する。

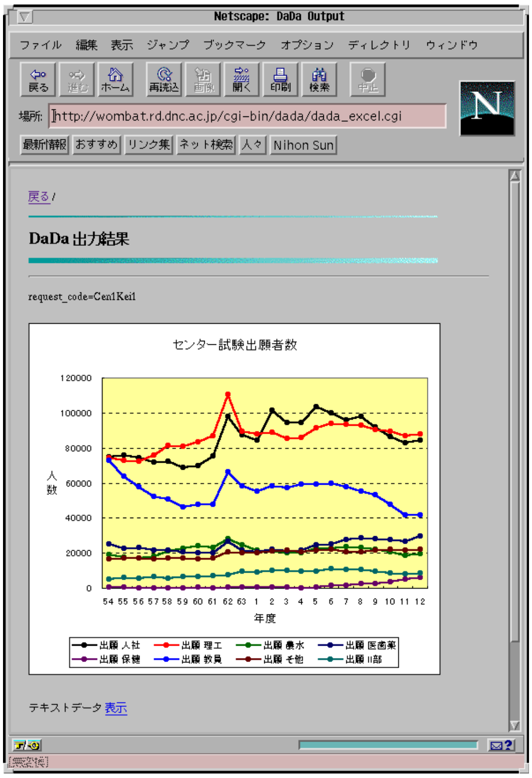
分類すべき項目は最大 2 つまで指定できる。指定しなくても構わない。複数を指定した場合は、その組み合わせごとに表示する。たとえば「学部系統別」と「性別」の 2 つを選択したなら、人文社会系を男女別に、理工系を男女別に、という風に表示する。

特に指定しなければ「学部系統別」に設定される。ここでは初期状態のまま、すなわち「学部系統別」のみを選択するものとする。

「検索条件クリア」で画面で指定した検索条件がクリアされる（初期状態に戻る）。

操作 3: 画面の「実行」ボタンをクリックする。

画面に excel の結果が図（gif 形式）で表示される（図 6）。



http://wombat.rd.dnc.ac.jp/cgi-bin/dada/dada_excel.cgi

DaDa 出力結果

request_code=Cen1Kei1

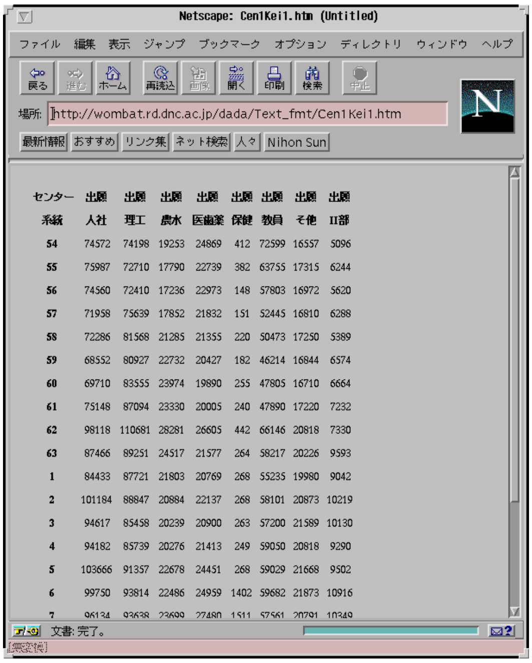
図 6: 出力結果

この結果は「センター試験の全出願者数」を、「学部系統別」に、すなわち「人社/理工/農水/医歯薬/保健/教員/その他/II 部」の 8 つの学部系統ごとに分類したものである。

図の下にある「テキストデータ表示」をクリックすると、その具体的な数値を得ることができる（図 7）。

検索開始画面では「操作方法ヘルプ」が利用できる（図 8）。単なる操作方法の説明だけでなく、操作上の注意事項や、「個別試験における出願学部の決定方法について」（付録 B に同文掲載）の説明が引けるようになっている。

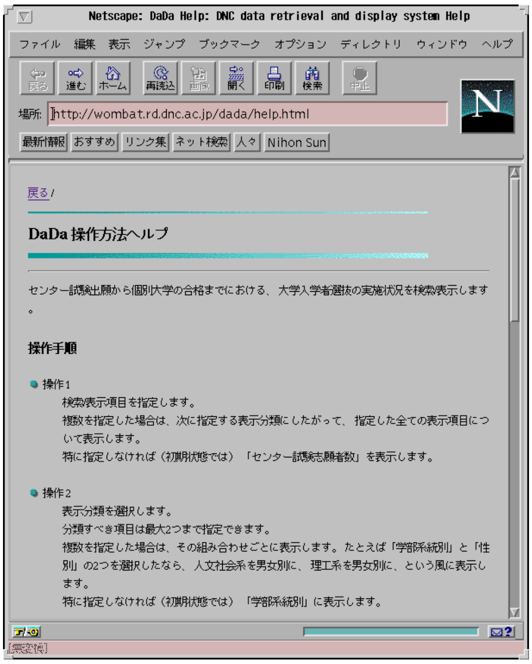
データ更新をすると自動的に「データ更新情報」が修正され、平成何年度までのデータが検索/表示できるかが検索開始画面で確認することができる（図 9）。



http://wombat.rd.dnc.ac.jp/dada/Text_fmt/Cen1Kei1.htm

センター	出願	出願	出願	出願	出願	出願	出願	出願
系統	**人社**	**理工**	**農水**	**医歯薬**	**保健**	**教員**	**その他**	**II部**
54	74572	74198	19253	24869	412	72599	16557	5096
55	75987	72710	17790	22739	382	63755	17315	6244
56	74560	72410	17236	22973	148	57803	16972	5620
57	71958	75639	17852	21832	151	52445	16810	6288
58	72286	81568	21285	21355	220	50473	17250	5389
59	68552	80927	22732	20427	182	46214	16844	6574
60	69710	83555	23974	19890	255	47805	16710	6664
61	75148	87094	23330	20005	240	47890	17220	7232
62	98118	110681	28281	26605	442	66146	20818	7330
63	87466	89251	24517	21577	264	58217	20226	9593
1	84433	87721	21803	20769	268	55235	19980	9042
2	101184	88847	20884	22137	268	58101	20873	10219
3	94617	85458	20239	20900	263	57200	21589	10130
4	94182	85739	20276	21413	249	59050	20818	9290
5	103666	91357	22678	24451	268	59029	21668	9502
6	99750	93814	22486	24959	1402	59682	21873	10916
7	96134	93638	23699	27480	1511	57561	20791	10349

図 7: テキストデータ表示



http://wombat.rd.dnc.ac.jp/dada/help.html

DaDa 操作方法ヘルプ

センター試験出願から個別大学の合格までにおける、大学入学者選抜の実施状況を検索表示します。

操作手順

- 操作 1
  検索表示項目を指定します。
  複数を指定した場合は、次に指定する表示分類にしたがって、指定した全ての表示項目について表示します。
  特に指定しなければ（初期状態では）「センター試験志願者数」を表示します。
- 操作 2
  表示分類を選択します。
  分類すべき項目は最大2つまで指定できます。
  複数を指定した場合は、その組み合わせごとに表示します。たとえば「学部系統別」と「性別」の2つを選択したなら、人文社会系を男女別に、理工系を男女別に、という風に表示します。
  特に指定しなければ（初期状態では）「学部系統別」に表示します。

図 8: ヘルプ画面

場所: http://wombat.rd.dnc.ac.jp/dada/Text_fmt/Update_date.htm

DaDa データ更新情報

平成 12 年度までの大学入学者選抜実施状況が表示できます。

データ更新日: 2001年04月23日

図 9: データ更新情報

B 個別試験における出願学部の決定方法について

出願学部を以下の規則に従い判定する。

- 1 校のみに出願しているときは、その大学・学部へ出願したものとする。
- 複数校出願した場合には、以下の優先順位に基づき出願校 1 校に限定する。
  優先順位 1 : 合格
  優先順位 2 : 辞退
  優先順位 3 : 不合格
  優先順位 4 : 未受験
- 辞退者とは合格したが辞退の申し出をした者、または前期日程試験に合格したが入学手続きをとらなかった者をいう。
- 同一の合否結果が生じた場合、前期日程 (A 日程) に出願した大学・学部へ出願したものとして取り扱う。

以下は 2 校受験した場合における、出願先の具体例を表にまとめたものである。

付表: 判定した出願先とその合否状況

出願 (1)	出願 (2)	判定した出願先
合格	不合格	出願 (1) 合格
不合格	合格	出願 (2) 合格
合格	辞退	出願 (1) 合格
辞退	不合格	出願 (1) 辞退
合格	合格	出願 (1) 合格
不合格	不合格	出願 (1) 不合格

C　国公立大学における学部系統一覧

【人文社会系】
文学部
外国語学部
国際文化学部
人間文化学部
人文学部
社会学部
社会福祉学部
行政社会学部
福祉社会学部
法学部
法経学部
経済学部
経営学部
商学部
商経学部
経営経済学部
経営情報学部
情報学部
社会情報学部
学芸学部
文教育学部
発達科学部
人間科学部
人間社会学部
総合人間学部
国際関係学部
国際学部
国際情報学部
人文社会科学部
法文学部
人文社会学部
文１系
文２系
文３系
文科一類
文科二類
文科三類
総合政策学部
地域政策学部

【理工系】
理学部
情報科学部
工学部
基礎工学部
鉱山学部
システム工学部
工学資源学部
工芸学部
芸術工学部
電気通信学部
情報工学部
理工学部
コンピュータ理工学部
総合理工学部
ソフトウェア情報学部
システム科学技術学部
環境理工学部
環境科学部
環境共生学部
生命理工学部
理１系
理２系
理３系
理科一類
理科二類
理科三類（北大）
第１類
第２類
第３類
第４類
第５類
第６類
第７類

【農水産系】
農学部
園芸学部
農学生命科学部
獣医学部
畜産学部
商船学部
水産学部
繊維学部
生物生産学部
生物資源学部
生物資源科学部

【医歯薬系】
医学部
理科三類（東大）
歯学部
医学専門学群
薬学部

【保健系】
看護学部
看護福祉学部
衛生看護学部
保健学部
保健医療学部
保健福祉学部
保健科学部
保健医療福祉学部

【教育系】
教育学部（教育学）
教育学部（教員養成）
体育学部
体育専門学群
文化教育学部
教育文化学部
教育人間科学部
教育福祉科学部
教育地域科学部

【その他の系】
家政学部
生活科学部
食品栄養科学部
栄養学部
生活環境学部
人間環境学部
環境人間学部
芸術学部
芸術専門学群
美術学部
美術工芸学部
デザイン学部
音楽学部
情報文化学部
文化学部
事業構想学部
地域科学部
第一学群
第二学群
第三学群
教養学部
文理学部
総合科学部
総合管理学部
総合情報学部
図書館情報学部

【I部】