University Statistics Survey and Utilization of Big Data

The Interfaced university DB with survey and service

UNESCO
OECD
ADEA
SEAMEO
AIPA
OEI
World Bank
CONTENTS

01 Introduction

02 Establishment University Statistics Survey System with University DB

03 Implication
Introduction 01
## I. Introduction of University Statistics Survey

### 1) History

<table>
<thead>
<tr>
<th>Year</th>
<th>History</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td># Dispersion of educational statistic survey items under “The Ministry of Education” • Several items were transferred to to specialized institutes</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td># Implemented the Public Disclosure of College Information, and increased the importance of university statistic data as a raw data</td>
<td></td>
</tr>
<tr>
<td>2012~</td>
<td># Started enhancement of statistical system construction using raw data of higher educational institutes • Establishment and improvement of the system from 2012 to present (2015) • Test operation for Big data construction by uploading DB at university directly (1H, 2015)</td>
<td></td>
</tr>
<tr>
<td>2015, 10</td>
<td># Enhancement of university statistics survey is now in progress</td>
<td></td>
</tr>
</tbody>
</table>
## I. Introduction of University Statistics Survey

### 2) The status of the survey

| Purpose | • To support policy and research through making the university statistics  
          • To provide nationwide statistical information services |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>• About 1,800 of all higher educational institutes including junior college, university and graduate school</td>
</tr>
<tr>
<td>Time point</td>
<td>• Twice a year (April 1st, October 1st)</td>
</tr>
<tr>
<td>Method</td>
<td>• Upload through standardized database system of higher educational institutes</td>
</tr>
<tr>
<td>Outcome</td>
<td>• Big data of university statistics and publications such as &quot;Statistical Yearbook of Education&quot;</td>
</tr>
</tbody>
</table>
| Utilization | • Government (The Ministry of Education) policy support data  
               • Support data of government project such as “Public Disclosure of College Information”  
               • Basic data of educational research & Nationwide web service |
I. Introduction of University Statistics Survey

3) The status of the system

- Basic research/Plan for enhanced survey with upgraded DB
- Development the enhanced survey system
- The third enhanced survey system is developing

*Started reorganizing the current higher education statistics since 2012 and have kept improving the third survey system*
I. Introduction of University Statistics Survey

3) The status of the system

- Summary of the service system construction
  • Budget : 3.5 millions (US dollars)
  • Construction Period :
    ✓ 2012 ~ 2014 (1st and 2nd were completed developing)
    ✓ 2015 (The third development is in progress)
    ✓ Expansion of DB contents: Curriculum etc.

- Contents
  • Individual DB about overall educational statistical survey
    ✓ Departments, students, faculties, curriculums etc.
I. Introduction of University Statistics Survey

4) Current status of Big Data construction

University

Enrolment DB
Faculty DB
Class DB

KEDI

Student DB
Faculty DB
Class DB

Customer

Web Service
Published -material
Service for Policy support
Service for research support

Submit

Raw Data

Total essential information

Collected total DB

Data Correction

Tuning

Analysis table

Analysis DB
I. Introduction of University Statistics Survey

4) Current status of Big Data construction

• Previous DB is converted to standardized DB by migration and it maintains a time series with new individual DB

• Maintain the newly collected individual DB by converting to a virtual DB which is a form of Data Mart.
Establishment of University Statistics Survey System with University DB
II. Establishment of University Statistics Survey System Using University Database

1) University Statistics Survey Using Big Data

- Collect all data related to student, faculty and class with standardized form and establish it as Big data

- The survey system and diagram (Continued)
II. Establishment of University Statistics Survey System Using University Database
II. Establishment of University Statistics Survey System Using University Database

1) University Statistics Survey Using Big Data

- Expected effect of Higher education statistics enhanced survey

  - Shorten the term of survey (6months → 1.5months)
  - Downsize the human resource for the survey (around 70% reduction)
  - Increase the reliability on data by collecting a raw data
    - Resolve the consistency issues of submitted data
    - Having a raw data increases data reliability
    - Increase data reliability by cross-validation analysis with external DB
II. Establishment of University Statistics Survey System Using University Database

1) University Statistics Survey Using Big Data

- Expected effect of Higher education statistics enhanced survey
  
  • Maximization the data linkage, utilization and service
    ✓ Data service focused on the limited tables
      → Provide a customized data (various index and dataset for statistical analysis)
    ✓ Linkage Analysis and utilization with external DB
II. Establishment of University Statistics Survey System Using University Database

2) Improvement

- Conversion of the input method from the table form into the raw data form
  - Every year, over 30 percent of departments change their name or curriculum
  - Since the department is the basic unit of the survey, it is necessary to organize department.

- Set a primary key for raw data
  - Absence of commonly used primary key (resident registration number) causes the duplication of survey contents
II. Establishment of University Statistics Survey System Using University Database

2) Improvement

- Need to decrease the risk of having any personal information
  - Organize personal information, protection and regulation
  - Systematic improvement: data encryption, firewall, etc.
III. Implication

1) The reliability is the most important

2) Necessary to ease survey burden

3) Use records (like file or DB, etc.) in each university as much as possible

4) Need a considerable time and finance for establishing Big data → required a step-by-step approach in the medium and long term
III. Implication

5) Various service plan consideration with Big data
   • Figuring out the demand for users should be the first
   • Progressing with not only one service plan, also various service plans depend on the situation or condition.

6) Professional institutes for leading the statistical survey, analysis and service
   • Develop various survey items, guidelines and standards according to a changing circumstance
   • Need a professional organization that supports to submit university DB
   • Need to train professional human resource for analyzing Big data
Thank you

Korean Educational Development Institute, Baumoe 1rd. 35, Umyeon-dong, Seocho-gu, Seoul 06762

+82-02-3460-0384  kediman@kedi.re.kr  http://hi.kedi.re.kr