

統合 ID 管理システム構想とその開発

佐野 洋

東京外国語大学 大学院総合国際学研究院

sano@tufs.ac.jp

概要：大学の情報センターは、情報システムの利用者管理が重要な課題となっている。例えば、個人ポータルの実現では、単に情報技術やプラットフォーム、アプリケーションの提供だけでなく、サービスを実質化するために、学務管理（学生管理）データとの連携が必要になるし、教職員への連絡では、学務事務関連（委員会名簿等）データとの連携が必要になる。情報センターは学内組織横断的な活動が不可欠である。さらに、利用者毎のサービス充実の実施だけでなく、災害時の大学構成員への連絡については、学内各組織との連携体制が必要且つ不可欠であることが認識されている。現在、本学情報センターは、全学の統合 ID 管理の枠組みを構想し、過去 1 年以上にわたり、学内構成員のデータ収集を行った。同時に、卒業生データを電子化するなど、統合 ID 管理に必要な資源収集や管理過程、組織上の関連性の調査を行ってきた。本稿は、これまでの活動結果とそれを基にパイロット開発を進めているシステムについて報告する。

1. はじめに

大学の社交関係の変化にともなう大学機能の進化要請に応じて、教育研究事業を支える情報基盤には、さらなる効率化や柔軟な対応力が求められている。

1.1. 大学の社交関係の進化

図 1 は従来の大学の社交関係を示す。それに対して図 2 は、現代のネットワーク時代のソーシャルコミュニティを示す。

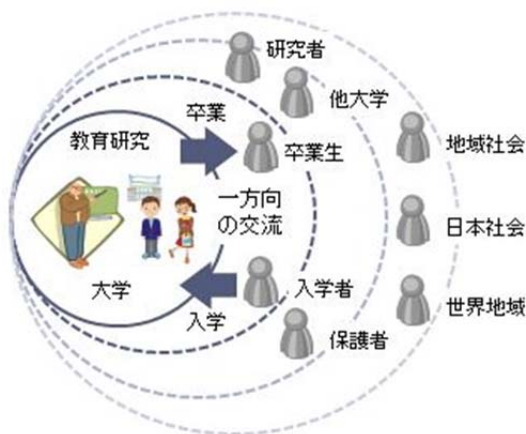


図 1

図で比較するように、教育研究組織である大学が社会と交わる有様は、従来の一方向の社交関係は、現在、双方向で且つ交流型の往交関係に変容している。

例えば、卒業生が大学等で勤務していれば、研究者コミュニティの構成員として本学と交流できる。これは研究組織としての大学の交流機能をも

って実現できていたことである。

ネットワークが社会基盤となり通信コストの劇的軽減が図られている現在、例えば、在学生の保護者が海外勤務していれば、世界地域コミュニティの構成員として本学と交流できるし、在学生が大学近隣に在住すれば地域社会コミュニティの構成員として本学と交流できる。

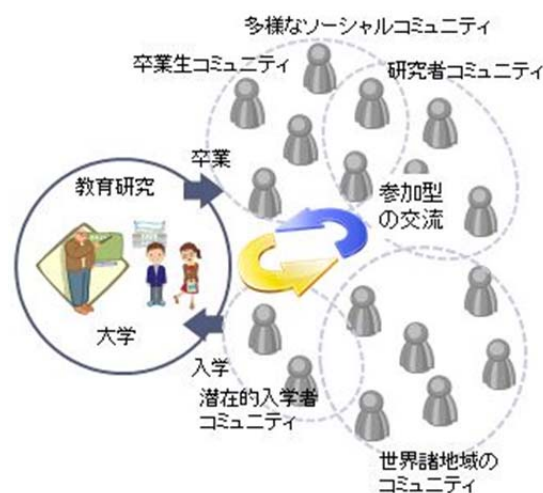


図 2

1.2. 構成員マネジメント

多様なソーシャルコミュニティに対応するための法人事業力と教育研究機動力の強化は火急に具体化するべき課題である。

具体化施策は、大学の構成員に対するマネジメント力の向上にある（協働のための絆基盤の構築）。大学法人に直接所属する構成員だけでなく、卒業生も一体として捉え、大学の社交力の強化（関係

者参加型の交流促進)を行う必要がある。

1.3. 統合 ID 管理システム構想

現在、本学情報センターは、全学の統合 ID 管理の枠組みを構想し、過去 1 年以上にわたり、学内構成員のデータ収集を行った。同時に、卒業生データを電子化するなど、統合 ID 管理に必要な資源収集や管理過程、組織上の関連性の調査を行ってきた。また、明治期(1900 年)以来の学位記台帳及び学籍簿から、卒業生の卒業時データを電子化した。同時に校友会(外語会)との協働についても会合を重ね、卒業生データの構築の準備を進めてきた。

本稿は、これまでの活動結果とそれを基にパイロット開発を進めている統合 ID 管理システムについて報告する。

2. 利用目的と整備データ

2.1. 利用目的・目標

教育研究に係るサービスの向上、法人競争力の向上を目指し、法人事業に携わる事業者、事業従事者、学生、卒業生を一意に特定できるようにすることを目的とする。

データ管理では、台帳データを維持管理する。データバックアップを兼ねる。データ更新では、ウェブデータベースアプリケーションを利用して、複数個所からデータを管理(追加、修正、削除等)ができるようにする。

2.2. 整備対象データ

表 1 には整備すべきデータ(教職員データ、在学生データ、卒業生データ)の現状をまとめた。

表 1

データ	関係組織(登録メンバー)	管理(状況)
教職員	人事、会計、各部署(非常勤教員、非常勤職員、派遣社員、業務委託等)	教職員番号 非正規の教職員や業務委託等については番号がない
在学生	教務(正規生、非正規生)、留学生担当(非正規生)、総務(市民聴講生などの非正規生)	学籍番号 入学年度や履修課程を表現
卒業生	学生担当、留学生	組織毎にデータを

	担当、校友会組織	管理
--	----------	----

教職員番号については、支払い管理と業務管理の区分が番号上されておらず、番号の意味づけが不十分である。

学籍番号は、履修制度に依存していて、履修制度が変わる番号の意味も変化する。従って、永続的にユニークな番号ではない。

卒業生については、利用している組織毎にデータ管理の仕方が違う。大学が管理する場合には、卒業年度と学籍番号を用い、校友会組織が管理する場合には、卒業年度と専攻名など在校時の履修情報を用いている。

整備すべきデータは、上記の 3 つのデータをまとめたもので、外語大組織に所属した時に生涯永続するユニークな ID を割り振り、その ID に、役割(ロール、時間情報付き)を付与して、個人のデータ管理を行うことで構築できる。課題を挙げる。

- データ収集と登録：素データの収集、素データのクレンジング、DB 及び DB アプリケーションの開発と運用
- データ管理の問題：管理責任者、データ管理責任者、データサービス管理責任者
- 協力体制の問題：学内協力体制、学外協力体制
- サービス：提供するサービス(開発する DB アプリケーションの設計に係る)

2.3. 整備データ

統合化された整備データの種類と利用イメージは以下である。

1. 台帳(電子データ)

既存の素データをクレンジングし、(ある程度の)名寄せを行う。マスターを作成した上でデータベースへの初期投入データとなる。その後、データベースを運用し、定期的にデータのスナップショットをとり、台帳データとして保存しておく。

2. データベース

マスターに基づき複数の管理者によってデータを維持管理する。ウェブデータベースアプリケーションによってデータベースにアクセスする。データベース及びデータベースアプリケーションは大学が中心になって維持管理

する。

3. 分析用データベース

事業推進のためのデータ資産（及びマスター妥当性の検討を行うことで 1、2 ヘフィードバックする）。

3. データ収集

3.1. 教職員データ

情報センターでは、過去 10 年にわたり申請されたアカウント情報を頼りに、教職員データの収集を行った。人事関係部署でヒアリングし、教職員情報で職能を確認した。ヒアリングの経緯で分かったことは、教員、事務職員などの雇用（職種）区分のデータ補足は、組織変更などを伴うと難しい。とくに、研究員や研修教員や非常勤事務職員などは業務期間の補足が難しいため、アカウント管理上もセキュリティの観点から維持・管理の必要性が高い。

本学の利用アカウントは約 600 で、過去のアカウントの発行記録から凡そ 1200 の教職員データを得た。情報システム運用上は、これらアカウントは削除したり、ロックしたりしている。

3.2. 在校生データ

いわゆる教務システムで管理されている学生は学部学生、大学院生である。研究生や科目等履修生だけでなく、例えば、交換留学生、研修留学生や教員研修留学生など、それぞれの学生数が少ないが学生カテゴリが多く、管理コストが高いデータがある。同時にこれらカテゴリの学生の管理部局は留学生課や留学生の教育組織であったりするので、学生状況の補足が困難であることが分かった。

3.3. 卒業生データ

大学が保有するデータは卒業時の卒業生データ（学位記、学籍簿）である。電子化されたのは 2000 年前後であり、それ以前は印刷データ（いわゆる台帳）がある。学位記台帳と学籍簿に記載のデータは以下の通りである。

データの 種類	媒体	管理者	データ内容
学位記 台帳	印刷	教務課	(学位) 番号、名前、 生年月日、教育課程 名、卒業年次

学籍簿	印刷、 電子デ ータ	教務課	入学年次、学籍番 号、教育課程名、性 別、連絡先(住所、電 話、メールなど)
-----	------------------	-----	---

印刷データとして、明治 33 年（1900 年）～平成 12 年（2000 年）までの学籍簿等がある。大学院は昭和 43 年（1968 年）～平成 13 年（2001 年）までである。印刷データの範囲で合計すると卒業生数は 32,280 名である。平成 13 年からは電子データに移行している。大学院は、平成 14 年から電子化されている。平成 20 年度現在で合計 39,447 名の卒業生があった。

印刷データは、電子化を行い、校友会等のデータを参照して、卒業後のデータとの名寄せ作業を行っている。

4. データ管理

4.1. データクレンジング

現在、データはエクセルを使って管理している。本節では、データクレンジングの経緯を説明する。

(1) 形式的データクレンジング手順

文字列は、基本的にはすべて全角文字に変更する。数字は半角だが文字列データに変換する。メール発達の有無や性別など 0/1 で表現されているデータは、データ項目毎に設定した記号に変換する。日付データは、1900 年以前のデータをチェックして、エクセル上で、統一した表示形式に変更する。

(2) 意味的なデータクレンジング手順

現在「未定」とか「空白」になっているデータは、各項目の記録の意味を考慮して「未定」や「不明」のデータを記録する。

文字列データ内の = 記号は、(a) 文字コード上の制約によって表示に使用しようとしている文字コードに表示したい文字が含まれていない場合に、(b)（紙の学位記を電子化したときにスキニングの結果として）判読不能文字の代替として使われているのだと推測できる。この記号は、(a)、(b) の 2 つの意味を持つものとして、そのまま残しておく。

僅かながら、1900 年以前のデータがある。数が少ないのでチェックして、正確な日付をメモとして記録している。これらは台帳で管理することとし、データベースに登録してからデータ修正を行う予定である。

なかには、不規則に存在している不明な表記（?とか…とか、メモ記述）がある。こうした場合には、データレコード全体をみて、判断する（修正手続きを一般化できない）。

4.2. 名寄せ

一個人に対して、一 ID を原則とする。ID は、「人工キー」とする。

卒業生データベースの名寄せでは、学部卒業生データと大学院卒業生データで重複データ（学部から大学院に進学した学生で同一の個人）があっても、名寄せでレコードを一行にまとめることはしない。

レコードを一行にまとめないことから、マスターの断面管理を行い、証跡データ（trail data）管理を実施することで、卒業生データを長期間にわたって維持管理する。ID が違うと別人として見なされ、トランザクションが発生したときの ID と、後に参照しようとしたときの ID が指す「個人」の実体（entity）が違っていることが起こる。この問題を避けるために、マスターキーに「適用開始年月日」を付加し、データ項目に「適用終了値月日」を追加する。

4.3. 統合 ID 管理システム

(1) システム機能要件

各既存システムとのデータ交換については以下の通りである。

- 学事、人事、校友会各システムと I/F が変化する
- 各既存システムからは ID 管理システムに必要な情報だけを取り込む機能を実装する
- ID 管理システムで発行するユニーク ID（生涯 ID）を各システムに戻す機能を実装する
インタフェースであるウェブ閲覧機能については以下である。
- 各部署の担当者が、適切なセキュリティー管理下で必要な情報をウェブ閲覧できる機能を実装する

データ登録・管理機能について以下に挙げる。

- 情報センターで受け付けた利用申請に基づき ID データ登録する機能を実現する
- 一括データ（CSV 等）で ID データ登録する機能

(2) 構成員テーブル

一個人に対して、一 ID を原則とする。

構成員コード(人工キー)

名前
読み(カナ)
アルファベット表記
生年月日
性別
所属開始
所属終了
現況
正副
参照構成員コード

ID は、「人工キー」とし、これは構成員テーブルとして実装し、各部局で参照するデータはこの ID をキーとして別テーブルとする。名寄せに必要な共通部分だけを設定し、管理することでデータ参

照時にも、限定された利用者に対し、制限されたアクセス権限を設定するようにする。

5. おわりに

本稿は、統合 ID 管理システムの構想と、実現のために実施してきた過去 1 年以上にわたる学内構成員のデータ収集、卒業生データの電子化等を説明した。それらデータ収集活動を基に開発を進めているパイロットシステムについて報告した。

統合 ID 管理システムは、大学の構成員に対するマネジメント力の向上の手段である。法人事業力を強化するもので、本学情報センターは、統合されたデータ管理の上に、ソーシャルネットワークシステムの活用による次世代コラボレーション機能強化を図る予定である。

また、教育支援の技術枠組みを利用して、卒業生との協働機能を備えたキャンパスライフ支援を行うこともできる。例えば、情報ボードの提供やコミュニケーション活性化などである。情報ボードの提供では、学務事務の業務効率の向上も目指している。

参考文献

- [1] 『きれいなデータの作り方』、日経 SYSTEMS、19 頁～37 頁、8 月号、2011
- [2] 林衛、「改訂 ER モデルによるデータベース設計技法」、ソフト・リサーチ・センター、2005
- [3] 佐藤正美、「データベース設計論—T 字形 ER」、ソフト・リサーチ・センター、2005