The Senate Information and Communications Technology Committee Annual Report
April 2011

This year the IT Committee focused on two issues:

- The current explosion in e-technology and its likely impact on education at Columbia
- The deficiencies of Data Governance policy, implementation, and dissemination at Columbia

The Committee presents reports on these two issues in this document.

Committee Members:

| | | |
|---|---|---|
| Fac. | Henry Spotnitz | P&S |
| Fac. | Itsik Pe'er *n-s* | SEAS |
| Fac. | Greg Bryan | A&S/NS |
| Fac. | Julia Hirschberg, CHAIR | SEAS |
| Fac. | Mark Cohen | BUS |
| Stu. | Vikey Suwanprutsachati | SCE |
| Stu. | Kenneth Durell | CC |
| Off. Res. | Hatim Diab *n-s* | |
| Libraries | Breck Witte | |
| Admin. | Candace Fleming *n-s* | |
| Admin. | Ellen Binder *n-s* | |
| Admin. staff | Norberto Govin *n-s* | |
| Alum. | Stephen Negron *n-s* | |

# E-Books and E-Readers in the Academic Setting - Information for the Columbia Community

### Senate IT Committee Report, April 2011

# Introduction

The last two years have brought tremendous advancements in digital access to otherwise to-be-printed information (*content*), through mobile electronic gadgets with varying screen sizes ranging from smart phones to those comparable in size to a book page. This document refers to both as *electronic books, or e-books.* For the first time in 500 years, the printed paper book has met a serious challenger. The consequences are particularly evident in the data-rich environment of a university. The Columbia Senate IT Committee has been engaged in a process of learning about this new technology, in order to allow the Columbia community to usher in the era of e-books in academic life.

## Diverse Needs

The Committee recognizes that different Schools have different needs with respect to eReading content and technologies. Diverse usage across campus roles, from student to administrator to faculty and faculty and students in different disciplines will have different interests in and requirements for eReading capabilities. The Committee believes that such needs can be well accommodated within the modular administrative structure of the university. Our goal in preparing this document is to bring the collected information about e-books to relevant stakeholders across schools and divisions of the university to facilitate potential action to welcome this expected and upcoming change in the academic experience.

## Content Is Key

When our investigations began late in 2009 (resulting in last year's survey on eReaders), we realized that a major obstacle to the widespread adoption of eReaders for classroom use was the lack of textbook material in electronic format. This situation has changed dramatically. Current widespread availability of content in many disciplines now facilitates a radical change in the student experience through student choice that is independent of the institution. We believe that preparation and accommodation on the part Columbia's schools can make this transition much easier for students and faculty alike. Specifically, schools should be taking steps to investigate whether content needed by their students is or can be made available in electronic format for the user's e-book devices. They should also consider whether they wish to partner with providers of hardware and content to make the transition smoother for all.

Below we describe some important considerations for schools investigating the transition to electronic content for their classes as well as individuals who are interested in the new technologies for their personal use.

# Standards and technologies

## e-Books/e-Documents Standards

Standards for e-book content are still evolving. E-Readers today typically allow users to import content from various input formats, even if it is not the preferred format for the device. A good example is the Portable Document Format (PDF), an open standard that is familiar to most computer users, which was developed long before e-Readers arrived on the scene; PDF fully describes a fixed-layout document (on creation). A newer free and open format, designed particularly for e-books, is the electronic publication (ePub) format. ePub is reflowable and text-centric, and therefore allows styling, such as dynamic word-wrap and font sizing. One of the main advantages of ePub is that it is device-independent; most handheld devices accept ePub format – with the notable exception of the Amazon Kindle.Proprietary vs. Open format capabilities is a major consideration for those considering eReader/eContent adoption. For example, is everything your students will need to read available for the eReaders they will use?

Other issues to consider when investigating eReaders and eContent providers include issues of Intellectual Property (IP) and Digital Rights Management (DRM). What rights will users have to use, annotate, and share annotations of the eContent they obtain – both for current course use and later consultation? Display capabilities obviously make a major difference in the eReader experience, as does enhanced functionality for annotation. The Form Factor of hardware devices is also important: Will the eReader be large enough for easy reading? Will it be light enough to carry to class?. Versatility, integration and connectivity: how are also important issues: How versatile is the device? How would How will it integrate with university systems and how connected is it? To investigate some of these questions, the Committee met with several leading vendors of eReader technology to compare platforms with regard to these questions.

# Platforms

The committee surveyed multiple e-book platforms for their suitability to the Columbia environment. We should note that the market is relatively new and evolving with many technologies, platforms and devices coming to market. Also, the design philosophy varies from from one device to the next. Both these circumstances make it difficult to compare platforms. However, below we describe two platforms, which we explored in some depth, as examples of available platforms. Each of these comprises of an e-Reader, but each also includes a distribution system for e-book content to go along with it. Such content systems may be, and in the examples below indeed are, cross-device systems, so adoption of one vendor's hardware does not preclude the use of the content by those with other devices.

## Apple's iPad

The iPad is a versatile general purpose 9.7" tablet computer. Since its release in April 2010, the iPad has redefined standards and expectations of devices between the smaller, handheld smartphone and the more versatile and conventional laptop. It is also used as an e-reader.

Many at Columbia are already experienced and avid iPad users, but only now is significant academic content emerging for iPads in the educational domain, prompting universities (such as Long Island University, locally) to give out iPads to incoming freshmen as a standard tool for student experience (and a hip attraction).

Apple offers iTunes U as a repository for academic material -- pdf docuents, videos and so on. Columbia as well as other leading universities uses iTunes U to share material. Apple offers an app store as a marketplace for third-party developers, providing an opportunity to diversify the options available to users. This model is expected to provide interfaces and applications for learning and other university functions.
**(Contact: Ronnie Peters, ronnie@peters@apple.com)**

## Kno

Kno is an e-book company focused on the student market. It offers an e-Reader which mimics a full-size textbook with one or two fold-in 14" screens. The e-Reader is a touchscreen tablet, allowing annotation and note-taking, with an interface for 3rd party developers. Leveraging their experience in textbook rentals, Kno offers content by most major publishers for purchase or rent.

Kno allows users to use their content platforms on other e-readers. The user's Kno account remembers the content that a user owns or rents. User-generated content, e.g., class notes or even comments and sticky notes within books, is attached to the account, rather than to the device, and is synchronized with a cloud-computing server.
**(Contact: Mike Hagerty, michaelh@kno.com)**

## XanEdu Publishing

XanEdu is a coursepack and custom textbook distributor with a specific focus on the e-Reader market. Among the many different institutions of higher education employing their services, a large number of business schools have utilized XanEdu to automate and digitize much of their case work; these are documents which usually consist of an onerous number of pages for many students to simply carry around in a full, physical paper form. XanEdu has ties to many publishers, but currently focuses most of their efforts on securing book or article-specific copyrights, customizing their product to faculty and schools through individual discussions and strategy sessions. XanEdu has a working relationship with Apple, which allows it to provide an iPad App for Higher Education free of charge to users. This app (an enhanced version of iBooks which uses ePub format) allowsnote-taking, highlighting, tracked changes, annotating, sharing highlights, etc,., both online and offline. XanEdu has yet to diversify their platform outreach; currently, the iPad device is the only mobile e-Reader for which they provide a distinct application. However, the company is testing other e-reading devices. Also, because the iPad App exports the coursepacks, shared contents and notes to PDF format only,smart, msmart phone users are hindered from viewingcontent,viewing the content as PDFs are unwieldy and cumbersome to read on current smart phone screens (e.g., iPhone, Android).

**(Contact: Tyler Steben, tsteben@xanedu.com)**

# Content

Vendors appreciate that the utility of their platform increases if users are able to leverage it to share and publish the content they generate. In an academic environment, this can represent an opportunity for teachers to increase their impact beyond the classroom.

# Integrating with LMSs

For many at Columbia, this spring marks a transition from the CourseWorks Learning Management System (LMS) to Sakai. Both LMSs are among the potential environments that vendors assume their platform will need to interact with, as part of the student's daily routine. The systems we have explored, as well as others, have developed interfaces that facilitate working with CourseWorks or Sakai.

# e-Book Technology Experiences at Columbia

Several Columbia schools are experiencing the e-book revolution currently, whether as an informal, grass-roots movement or through a planned trial. To aid schools that might be interested in providing eReader platforms or content to their students, we present some current instances of eReader adoption and trials at Columbia. We also note that the Columbia Libraries have been in the forefront in recognizing the importance of supporting eReader technologies and have considerably increased their digital collections over the past years.

## Business School Trial

Janet Horan, Vice Dean of the Columbia Business School, informs us that the school will test the use of XanEdu this summer in two courses that are part of the Executive MBA program. This is potentially a major step forward in taking advantage of these emerging E education technologies and techniques.

**(Contact: Mark Cohen, mac2218@columbia.edu)**

## iPad Use at CUMC

At CUMC, the iPad appears to be the first reader for the Electronic Medical Record to be successfully integrated into the daily workflow of patient management. The Department of Surgery has envisioned use of mobile computers on surgical rounds for more than ten years, but this goal has been elusive until now. Factors contributing to recent success:

- Hospital-wide systems implemented in the past two years include electronic x-ray viewers (GE Web), ECG viewers, and EMRs for inpatients (Eclipsys XA) and outpatients (Allscripts Touchworks or "CROWN"). The EMRs allow viewing of lab results, order entry, and entry of clinical notes for patient management.
- Citrix emulation was used for the EMRs, allowing access by both PCs and Macs.

- A recent upgrade by Citrix to their free emulator for the iPad provided fully functional access for the iPad to XA and to CROWN (and via CROWN to GE Web).
- The long battery life (more than 10 hours), portability, wireless connectivity, and convenience of the iPad operating system (touch interface and instant on) make it extremely useful for work and contrast dramatically with older technologies.
- Security of the iPad appears adequate, provided that a long password is implemented and remote wipe capabilities are implemented in the event of theft, although this issue will require more investigation.

Unresolved problems include how quickly the devices will be lost or stolen and how to convince users not to store Protected Health Information (images, patient lists) on the devices.

The iPad is currently being used daily on the Critical Care Service, led by Dr. Tracey Arnell. During patient work rounds, the device is used to view patient records, labs, and x-rays. It is used to write notes and to order medications and lab tests. An MD billing module under development by CROWN is currently in beta testing, but this has not yet been integrated into daily work rounds. Responding to Dr. Arnell's success, other division chiefs in Surgery have ordered iPads for similar purposes. Several house officers are using personal iPads for clinical work. Note that the iPad is considered a beta device by NYP and CUMC and is not officially supported. However, there is a strong impression that the iPad is the first of a series of mobile tablets that will create immense commercial pressure to continue to improve the technology of mobile computing.

**(Contact: Henry Spotnitz, hms2@columbia.edu)**

## Columbia Libraries

In the Columbia Libraries, it has long been recognized that modes of scholarly publication, research and learning processes and their outputs are rapidly evolving. Large parts of the collections that libraries have built over the past century or more are becoming available digitally and will need to be reacquired and preserved. In some areas print publication will be further marginalized and may disappear completely. E-books are gaining greater acceptance as an alternative to print acquisition, media such as print news sources and music CDs are being replaced by online formats following new business models, and non-commercial scholarly content may be disseminated only on the web. These trends have contributed to an accelerated transition to the adoption of electronic formats as content becomes available. A good starting point for learning about Columbia's online collections and services is the Libraries' web site: http://library.columbia.edu/index.html and for more detailed information about electronic resources: http://library.columbia.edu/eresources.html. Columbia University Libraries is committed to preserving these resources, to building systems that facilitate their discovery and use, and to assuring that they are accessible from any location and across the wide variety of computing devices our users may have at hand.

**(Contact: Breck Witte, witte@columbia.edu)**

# Recommendations

One way to deal with the emergence of e-books on campus is not to deal with it at all. At Columbia, this would mean that CUIT would set up applications to support e-readers on the network just as other devices are supported, but will not leverage them in particular. However, there are several ways that Columbia and individual schools could take advantage of and support a change that the Committee believes is surely coming.

a. School and Department Level:
   Individual schools or departments, with help from CUIT, could implement mechanisms for publishing and distributing electronic materials for courses. These efforts could be preceded by trials in designated classes to understand how best to support a larger effort. The Committee's 2010 survey clearly showed that different schools and departments have different attitudes toward the adoption of e-readers. The independence of units across Columbia facilitates making decisions for different levels of adoption in different schools, and the e-book platforms are flexible to facilitate heavy use when the user is in one context vs. only passive adoption in another (for instance, when a student takes classes across schools). A broader level of adoption might then involve declaring a broader change to an e-book format (e.g. for an incoming class to a program).
b. Single Class Pilots
   At the level of a single class, the usefulness of e-books can be explored and evaluated in terms of enriching interactivity of classroom experience. Depending on the class, and the degree to which the class material is available in digital form, this may involve relatively little effort on the part of faculty or students.

**Committee Recommendation:**

**The Senate IT Committee strongly recommends that each School appoint a particular individual within the school administration to manage the increasing use of e-book technologies in the School and to explore opportunities to expand -- and potentially institutionalize -- the use of e-book technologies in the classroom. This individual should consider the following possible steps toward this exploration:**

- **Encourage faculty to include electronic versions of course data where available and advertise this availability to their students.**
- **Identify whether textbooks used in the School are available in e-version and assess pricing differences.**
- **Contact vendors of e-book devices and e-content distribution aggregators to familiarize themselves with what content is available and what devices may be most suitable for their school.**
- **Organize e-book vendor fairs for students and faculty.**
- **Pilot programs with individual faculty or courses.**
- **Serve as the e-book/e-content point of contact for vendors, faculty, and students and advertise this prominently on the School website.**

Columbia University

# Senate IT Committee on Data Governance
# April 22, 2011

March 2011

# Policies, Procedures, and Training Subcommittee

## Charter

Policies, Procedures, and Training Subgroup: Columbia University has a large number of well-developed and well-defined policies that govern data management and security. Nevertheless, it quickly became apparent, that many members of the Columbia University community may not realize how important these policies are to their studies or work at Columbia or may not even know that these policies exist. While it is essential that central administration is well informed about data governance policies, our goal was to recommend strategies to educate and inform the more decentralized parts of the university, mainly those at the school and department level.

## Members

Kenny Durell, Student, Columbia College

Candace Fleming, Vice President, Information Technology, CUIT

Norberto Govin, Data Manager/Analyst, Office of the Provost for Institutional Research

James Lindner, Assistant Vice President, Human Resources

Stephen Negron, Alumnus, President of SANTech Network Consulting, Inc.

Rose Razaghian (Chair), Director, Planning and Analysis, Arts and Sciences

Henry Spotnitz, M.D., George H. Humphreys, II Professor of Surgery, CUMC

# Policies, Procedures, and Training

## Recommendations

1. Educate Columbia University Community about Key Security and Data Governance Policies
   - Focus on School- and Department-level: Students, Faculty, Staff
   - Identify key security and data governance policies for each group
   - Create template posters for distribution by target audience
   - Implementation charged to CUIT and Communications
   - Group to continue to serve as advisory body
2. Initiate a long-term strategy to Continue Education and Awareness
   - IT Committee to explore in more detail alternative venues for education and awareness (e.g. certification, file sharing alternatives, etc.)

# Data Quality and Data Management Subcommittee

## Charter

Data Quality/Data Management – the goal of this group is to evaluate the types of obstacles that exist for employees to effectively use University data and to formulate strategies to overcome such obstacles. Such strategies should consider data management best practices and what other higher education and corporate institutions successfully implement. Such strategies should be pragmatic and include short-term benefits while delivering tools and processes that will enable greater long-term business value.

## Members

Jennifer Caplan, Associate Registrar, Registration and Financial Services

Lucy Drotning, Associate Provost, Office of Planning and Institutional Research

Ron Forino, Director, Enterprise Reporting, CUIT

Norberto Govin, Data Manager/Analyst, Office of the Provost for Institutional Research

Itsik Pe'er, Assistant Professor, Computer Science

Sue Spencer, Assistant Director, Human Resources

# Data Quality and Data Management

## Recommendations

1. Implement Data Stewardship Program
   - Identify and confirm Data Stewards and Data Architect(s)
   - Define roles and responsibilities
   - Collect data definitions, create platform that provides access to data inventory and definitions. Include security “category” within definition process.
   - Meet periodically to report on progress
2. Implement Data Dictionary (metadata collection tool)
   - Tool will facilitate the collection and dissemination of data definitions
   - Data Architect to inventory data found in reporting databases
   - Data Stewards to provide field level definitions
3. Following the implementation of recommendations 1 and 2, implement Data Quality initiative

# Privacy, Compliance, and Security Subcommittee

## Charter

This sub-committee will start with an enterprise scope, but as the effort develops it should be limited to specific types of data (Student Data). The effort should include researching and recommending technologies or methodologies that will help Columbia University locate and protect sensitive data. The subcommittee will try to determine how current data security policies and controls are managed and implemented across the university.

## Members

Medha Bhalodkar, AVP, Chief Information Security Officer

Norberto Govin, Data Manager/Analyst, Office of the Provost for Institutional Research

Rebecca Hirade, Associate Dean, Graduate School of Arts and Sciences

Bashir Khan, Senior Director, Financial Information Systems, CUMC

Tim Qin, Student, Columbia Engineering

Paul Brenner, Student, Columbia Engineering

Vikey Suwanprutsachati, Student, Continuing Education

## Recommendations

1. Short-term, to expand the Appendix A of the Data Classification Policy with other representative examples of data and their sensitivity levels. This will help our various data players to have a better understanding of what data are highly sensitive and what data are not.
2. Long-term, at the time when the metadata for Columbia University data systems are created, the data security classification must be added to it. This classification should be done by the data owners or data stewards on a solid understanding of University and government policies (Red Flags, FERPA, HIPAA, AARA, etc.). All internal analyses and reports must include the sensitivity level of the data in the report header.

# Senate IT Committee Overall Recommendation

## Creation of a Permanent Data Governance Committee

The goals of this committee will be:

- To develop data definitions and policies which are consistent, uniform, and accurate for reporting and analyses (Master Metadata Repository).
- To provide a framework which ensures data are captured accurately and consistently to maximize clarity of common usage, allowing management to make informed, shared, and timely decisions (Data Quality Management).
- To identify roles and responsibilities among key data players and to establish uniform data management standards and guidelines for them to use (Management of Roles and Responsibilities).
- To make sure the new data repository will include the security sensitivity classification system adopted by the technology security office. This will allow application owners and data custodians to take the necessary steps to secure our information systems (Data Security Compliance).
- To advise senior managers of which steps could be taken to increase awareness of the importance of complying with Columbia University information security guidelines and other data governance-related policies and procedures.

Senate Data Governance Subcommittee

Policies, Procedures, and Training Subgroup Report and Recommendations

Members:

Kenny Durell, Student, Columbia College
Candace Fleming, Vice President, Information Technology, CUIT
Norberto Govin, Data Manager/Analyst, Office of the Provost for Institutional Research
James Lindner, Assistant Vice President, Human Resources
Stephen Negron, Alumnus, President of SANTech Network Consulting, Inc.
Rose Razaghian (Chair), Director, Planning and Analysis, Arts and Sciences
Henry Spotnitz, M.D., George H. Humphreys, II Professor of Surgery, CUMC

Charter:

Policies, Procedures, and Training Subgroup: Columbia University has a large number of well-developed and well-defined policies that govern data management and security. Nevertheless, it quickly became apparent, that many members of the Columbia University community may not realize how important these policies are to their studies or work at Columbia or may not even know that these policies exist. While it is essential that central administration is well informed about data governance policies, our goal was to recommend strategies to educate and inform the more decentralized parts of the university, mainly those at the school and department level.

Report:

Columbia University has a large number of well-developed and well-defined policies that govern data management and security. Nevertheless, it quickly became apparent, that many members of the Columbia University community may not realize how important these policies are to their studies or work at Columbia or may not even know that these policies exist. While it is essential that central administration (such as finance, HR, IT) are well informed about data governance policies, our focus here is on the more decentralized parts of the university, mainly those at the school- and department-level.

We make two sets of recommendations: Those that are short-term which can be quickly and cheaply implemented and those that are long-term looking to the future of educating the Columbia community.

For the short-term, we recommend that students, faculty, researchers, and administrators be educated about the most important policies on data governance, which mostly center on information security. We reviewed the data governance policies that currently exist and from this pool selected the policies that Columbia individuals at absolute minimum need to be aware of. We have summarized these policies and drafted a small poster that could become a template for schools and departments as well as a starting point for the specific policies that may be most relevant to their divisions as they consider how to best educate their members on an ongoing basis. Each template also includes a keyword for each corresponding policy that, when typed into the Columbia University search engine, will come up with the official University policy first and other search results subsequently. We have constructed two templates, one for students and one for faculty or administrators. Of course, each

school and/or department may want to customize these templates for their constituents. (Please see Appendix 1 and 2)

We recommend that CUIT and the Communications group within Student Services be responsible for the implementation of this recommendation, including vetting the templates provided here, publishing them, programming the intended search results from the specified keywords, and developing and carrying out a strategy for dissemination including online mechanisms geared toward specific audiences. For example, "Tip of the Day" prompts could be inserted in the external login page to Outlook, SSOL, or CourseWorks, among others. We have also laid out some of the possible ways in which students, faculty, and staff can be educated about university data governance policies on an on-going basis. (Please see Appendix 3 for possible avenues of dissemination.) Further, we recommend that the currently created group on Policies, Procedures, and Training be continued as an advisory board for CUIT and the Communications group. The group will serve as a sounding board for CUIT and the Communications group, be updated on a bi-annual basis by CUIT and the Communications group on actions taken to implement the recommendations, and report progress on a bi-annual basis to the Senate Information Technology committee.

In the long-term, we recommend that the Senate Information Technology committee continue their work on how the Columbia community can be further educated about data governance policies. Among the alternatives, we recommend that the committee consider a certification program that requires all or some groups to demonstrate their knowledge of data governance policies via an electronic questionnaire and to explore file sharing technology such as Accellion file transfer.

At this point, the central goal is for <u>all</u> Columbia community members to be aware of the University's most important data governance policies. Compliance with policies will reduce the impact and number of security breaches, as much as possible, and mitigate their risk.

*Appendix 1: TEMPLATE FOR STUDENTS*

## HAVE YOU EVER WONDERED WHAT WOULD HAPPEN IF SOMEONE GOT A HOLD OF YOUR DATA?

## DON'T WONDER: PROTECT YOUR PROPERTY AND INFORMATION

*For More Information Type this Keyword into the CU Search Engine*

- **PROTECT YOUR COMPUTER, LAPTOP, AND HAND-HELD DEVICES:**

  **Strong Passwords**

  - Don't let your **Password** get hacked
    - Use strong passwords that are:
      - 8 characters or longer
      - Combination of UPPER and lower case letters
      - Include numbers, spaces or special characters whenever possible (%,?,@, etc.)
    - Do NOT share your password
  - Protect access: Enable **Screensavers** — **Screensavers**
  - Guard against Hackers: Install a build-in **Firewall** — **Firewall**
  - Protect against Computer **Viruses** with Anti-Virus Software: — **Anti-Virus**
    - Symantec Antivirus is FREE for Columbia students, faculty, staff
  - Protect Information: Protect files with **Encryption** Software — **Encryption**
    - *Guardian Edge* is the recommended software
  - Don't lose your work: Make **Backups** frequently — **Backups**

- **PROTECT YOUR DATA:** — **File Sharing**
  - Do NOT use peer-to-peer file sharing software (e.g. emule or utorrent)
  - Do NOT post sensitive information on the web (e.g. google docs or groups)
  - Only if absolutely necessary email files:
    - Password protect files with sensitive or confidential information
    - Communicate the password separately

- **SANITIZE YOUR COMPUTER** — **DBAN**
  - If recycling or discarding an old computer, sanitize the hard drive by using DBAN

- **REPORT A SECURITY BREACH ASAP**
    - Contact: security@columbia.edu

*Appendix 2: TEMPLATE FOR FACULTY or ADMINISTRATORS*

## HAVE YOU EVER WONDERED WHAT WOULD HAPPEN IF SOMEONE GOT A HOLD OF YOUR DATA?

## DON'T WONDER: PROTECT YOUR PROPERTY AND INFORMATION

*For More Information Type this Keyword into the CU Search Engine*

- **PROTECT YOUR COMPUTER, LAPTOP, AND HAND-HELD DEVICES:**
  - Don't let your **Password** get hacked — **Strong Passwords**
    - Use strong passwords that are:
      - 8 characters or longer
      - Combination of UPPER and lower case letters
      - Include numbers, spaces or special characters whenever possible (%,?,@, etc.)
    - Do NOT share your password
  - Protect access: Enable **Screensavers** — **Screensavers**
  - Guard against Hackers: Install a build-in **Firewall** — **Firewall**
  - Protect against Computer **Viruses** with Anti-Virus Software: — **Anti-Virus**
    - Symantec Antivirus is FREE for Columbia students, faculty, staff
  - Protect Information: Protect files with **Encryption** Software — **Encryption**
    - *Guardian Edge* is the recommended software
  - Don't lose your work: Make **Backups** frequently — **Backups**
    - Use Columbia backup software
    - Make backups on shared drives
- **KNOW IF YOU WORK WITH OR STORE CONFIDENTIAL DATA:** — **Data Classification**
  - The most confidential and sensitive data includes:
    - Social Security, Drivers License, and Credit Card Numbers
    - Student records (academic, financial, personal, etc.)
    - Medical records
    - Financial, personnel, payroll records
- **PROTECT YOUR DATA:** — **File Sharing**
  - Store and share sensitive data on a shared server
  - Do NOT store files on the hard drive of your desktop or laptop whether at the office or at home
    - If necessary to store on the hard drive, password protect files
  - Do NOT use peer-to-peer file sharing software (e.g. emule or utorrent)
  - Do NOT post sensitive information on the web (e.g. google docs or groups)
  - Only if absolutely necessary email files:
    - Password protect files with sensitive or confidential information
    - Communicate the password separately
- **SANITIZE YOUR COMPUTER** — **DBAN**
  - If recycling or discarding an old computer, sanitize the hard drive by using DBAN

- **REPORT A SECURITY BREACH ASAP**
  - Contact: security@columbia.edu

Appendix 3

POSSIBLE AVENUES OF DISSEMINATION OF DATA GOVERNANCE POLICIES, BY GROUP

I. Students
   a. School orientation sessions
   b. Posters in computer labs (next to every computer or printing station, Student Services, Lerner, lobbies in the schools)
   c. School Websites
   d. Course Management Portal, SSOL

II. Faculty/Officers of Research
   a. HR packet
   b. Faculty orientation (annual)
   c. New chair orientation (annual)
   d. Dean meetings with chairs (annual)
   e. Annually information sent to deans and department chairs for review of current policies and introduction of new policies
   f. Website

III. Administrators/Department Administrators
   a. Hiring orientation
   b. Training sessions
   c. Academic Department Administrator meeting (annual)
   d. Annually information shared with Managers for review of current policies and introduction of new policies
   e. Posters that can be posted in centrally located offices
   f. Website

Senate Data Governance Subcommittee

Data Quality and Data Management Subgroup Report and Recommendations

Members:

Jennifer Caplan, Associate Registrar, Registration & Financial Services
Lucy Drotning, Associate Provost, Office of Planning and Institutional Research
Ron Forino, Director, Enterprise Reporting, CUIT
Norberto Govin, Data Manager/Analyst, Office of the Provost for Institutional Research
Itsik Pe'er, Assistant Professor, Computer Science
Sue Spencer, Assistant Director, Human Resources

Chapter:

Data Quality/Data Management – the goal of this group is to evaluate the types of obstacles that exist for employees to effectively use University data and to formulate strategies to overcome such obstacles. Such strategies should consider data management best practices and what other higher education and corporate institutions successfully implement. Such strategies should be pragmatic and include short-term benefits while delivering tools and processes that will enable greater long-term business value.

Report:

Columbia University has implemented a number of reporting systems and databases being supported at the enterprise, department, and the personal desktop level. It quickly became apparent that the information assets of the University are not leveraged as widely or as efficiently as possible. Data contained in both procured 3rd party software and internal custom built systems are not widely known, nor are they well understood except to the small group of trained and experienced information consumers who use them. This lack of understanding has resulted in lower that acceptable confidence levels in data accuracy for those developing reports and queries and for those who use them. Furthermore, the way data is organized within reporting database is not optimal. Many are overly simplistic in their data architecture, characterized as being "data silos" that lack integration, performance optimization, and ease of use. There is a common misconception that best-of-breed reporting tools can overcome the shortcomings of un-integrated data. In fact, well architected and well documented reporting databases will yield higher performance from report developers and consumers, decrease the risks associated with reporting inaccurate data, enable improved data analysis, and minimize data anomalies – regardless of the reporting tool. To address these opportunities, we make the following recommendations.

The quickest return-on-investment will come from implementing the tools, processes, and defining new roles to identify, organize and define CU's data assets. Subsequently, our first recommendation is to implement a data stewardship program. Such a program will identify roles and responsibilities of CU groups and individuals who would manage, define, and govern specific areas of data (i.e. general ledger or student registration – see appendix). Currently, these individuals already exist in functional areas; this program would more clearly identify their responsibilities to consumers of

their data. In such a Data Stewardship program, the Data Steward would be responsible to define their data and to be accountable for the data's accuracy and completeness. CUIT will be responsible for the role of the Data Architect, who would support a Data Dictionary or Metadata Repository, and who would create baseline information from existing and new reporting databases. CUIT's responsibility would cover enterprise databases and to assist in the collection of metadata from departmental database.

The second recommendation would be to provide the Data Stewards with a tool to collect and maintain a data dictionary for all sharable reporting databases. A data inventory would be available to all potential report developers and consumers, and would include information on a field-by-field basis providing definitions, formulas, allowable values/ranges, field size and type, etc. It would also include information about each field's security "classification" as defined by the Data Steward in conjunction with the Security Office. The repository tool would also include the ability for the Data Architect to design new reporting databases using best practice techniques (i.e. data models, data mart design techniques).

Following the implementation of these recommendations, as a subsequent initiative, we are also proposing that steps be taken to improve the information quality of the data being collected within CU's reporting databases. This initiative would be further studied and specific recommendations made to improve the quality and usability of data within CU's 3rd party and custom built reporting applications. Recommendations will address how to be more "proactive" in preventing data anomalies, and how to build better "reactive" processes to tracking and remedy data issues when they do occur.

## Appendix – Proposed Data Stewards

| Data Subject Area | Data Steward/Data Manager |
|---|---|
| DATA EXECUTIVE – EVP FOR FINANCE | |
| Budget (COB) | Nancy Johnson |
| | Jill Goldstein |
| Financial (FAS, FFE) | Ed Hamilton |
| Labor (LDS) | Frank Iovino |
| Procurement (APCAR) | Terry Park |
| | Anthony Cerbins - Interim |
| DATA EXECUTIVE – EVP FOR ADMINISTRATIVE SERVICES | |
| Human Resources (PAC) | Jim Lindner |
| | Paul Meck |
| Student Services (SIS, PF) | Laurie Schaffler |
| | Barry Kane - Registrar (Academic Records) |
| | Anthea Jefferies - Student Financial Services |
| | Mercy Goodnow-Smith - Financial Aid |
| DATA EXECUTIVE – EVP FOR RESEARCH | |
| Sponsored Projects Admin | Carol Tyko |

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

# IT SENATE DATA GOVERNANCE SUBCOMMITTEE

## Privacy, Compliance and Security Subgroup Report and Recommendations

Members:

Medha Bhalodkar, AVP, Chief Information Security Officer
Norberto Govin, Data Analyst, Provost Office
Rebecca Hirade, Associate Dean, Graduate School of Arts and Sciences
Bashir Khan, Senior Director, Financial Information Systems, CUMC
Tim Qin, Student, Columbia Engineering
Paul Brenner, Student, Columbia Engineering
Vikey Suwanprutsachati, Student, Continuing Education

Charter:

This sub-committee will start with an enterprise scope, but as the effort develops it should be limited to specific types of data (Student Data). The effort should include researching and recommending technologies or methodologies that will help Columbia University locate and protect sensitive data. The subcommittee will try to determine how current data security policies and controls are managed and implemented across the university.

Report:

Columbia University is committed to respecting and protecting the privacy of its students, faculty and staff, as well as to protecting the confidentiality of the information used to carry out the University's academic and research mission. Columbia University has identified three categories of data for the purpose of determining who is allowed to access the information and what security precautions must be taken to protect the information against unauthorized access. The categories are listed below:

1. Category HS – Highest Sensitivity (Confidential / Sensitive Data)
2. Category MS – Moderate Sensitivity (Internal / Official Use Only Data)
3. Category NS – Non-Sensitive (Public Data)

The group makes two recommendations, one which can be quickly implemented and one which is long term and requires University-wide Master Metadata to be put in place first. Metadata is defined in the Merriam-Webster dictionary as data that provides information about other data.

For the short term, we recommend that the Appendix A of the Data Classification Policy be expanded with other examples of data and their sensitivity levels. This document provides the list of representative examples of data that were classified using Columbia University data classification guidelines during the meetings of this group. This will help our different data players to have a better understanding of what data are highly sensitive and what data are not.

In the long term, we are recommending that at the time when the metadata for Columbia University data systems are created, that the data security classification be added to it. This can be done at the table or the attribute level. This classification should be done by the data owners or data stewards on a solid understanding of University and government policies (Red Flags, FERPA, HIPAA, AARA, etc.). Also, all internal analyses and reports must include the sensitivity level of the data in the report header. Data managers and reporting analysts should be able to find the sensitivity of the data by searching the master metadata repository.

See below for specific examples of data classification; the list is not all-inclusive:

1. **HS – Highest Sensitivity (Confidential / Sensitive Data)**
   - Credit card numbers
   - Bank account numbers
   - ABA routing numbers
   - Date of birth
   - Place of birth
   - Social security numbers
   - Driver's license numbers

2. **MS - Moderate Sensitive (Internal / Official Use Only Data)**
   - Student tuition bills
   - Payment history
   - Financial aid / grant information / loans
   - Athletics or department recruiting information
   - Class lists or enrollment information
   - Grades / Transcripts
   - Teaching assistant reviews
   - Permanent mailing address
   - Residence assignment
   - Degree(s) awarded and date(s)
   - Major(s), minor(s)

3. **NS – Non-Sensitive (Public Data)**
   - Student name
   - Directory address and phone number
   - Columbia university e-mail
   - University network ID
   - School and department of enrollment

# COLUMBIA UNIVERSITY

# DATA GOVERNANCE COMMITTEE (DGC) CHARTER

Data governance is a framework of policies, procedures, and standards used to establish the effective use and collection of a valuable asset, “University data.”

It is the recommendation of the Senate Information Technology Committee that a University-wide Data Governance Committee be established. This cross-functional committee will serve as an advisory body to resolve data issues and establish policies and definitions for the proper handling and interpretation of University data. These definitions will be stored in a data repository which will contain details of each element from technical, functional and security perspectives. The committee will consider data governance issues that have broad data and reporting implications, and will communicate their recommendations to senior management.

More specifically, the goals of this committee will be:

1. To develop data definitions and policies which are consistent, uniform, and accurate for reporting and analyses (creation and publication of a Master Metadata Repository).

2. To provide a framework which ensures data are captured accurately and consistently to maximize clarity of common usage, allowing management to make informed, shared, and timely decisions (Data Quality Management).

3. To identify roles and responsibilities among key data players and to establish uniform data management standards and guidelines for them to use (Management of Roles and Responsibilities).

4. To make sure the new data repository will include the security sensitivity classification system adopted by the technology security office. This will allow application owners and data custodians to take the necessary steps to secure our information systems (Data Security Compliance).

5. To advise senior managers of which steps could be taken to increase awareness of the importance of complying with Columbia University information security guidelines and other data governance-related policies and procedures.

The Senate IT Committee believes that leadership from both the academic and administrative sides of the University should sponsor such a committee in partnership with the University Senate.

Sponsors:

- Provost
- Senior Executive Vice President (SEVP)
- Chair, Senate Information Technology Committee

The membership of the committee should include representatives from a wide range of offices across the University that are responsible for collecting, maintaining, and reporting on the University's data. We propose that the committee should meet monthly.

We suggest that members could be drawn from the following groups:

1. Office of Management and Budget
2. Controller
3. Treasurer
4. Procurement
5. Human Resources
6. CUIT Enterprise Reporting
7. Student Services
8. Arts and Sciences
9. Office of Planning and Institutional Research
10. Medical Center Information Systems
11. University Registrar
12. IT Senate Committee Members
13. CUIT Management
14. Information Security
15. Vice Provost for Academic Administration
16. Student Senators
17. Representatives from the Schools
18. Columbia Libraries
19. Faculty Members

The Senate IT Committee believes that offices representing both the Provost and the SEVP should spearhead the organization and management of such a Data Governance Committee. Within the Provost's Office, the Office of Planning and Institutional

Research (OPIR) could take on this role and the Office of Financial Systems could participate on behalf of the SEVP.