Class Notes

Class 1 — Overview, Course Rules

Overview

- Topic: using computer and network technology to help run businesses and other organizations
- Won’t focus especially on “managers”
- Will combine “Top-down” descriptive learning (the RTP book) with “bottom-up” learning by example (Microsoft Access and GB book)

Rules and Procedures – see the syllabus and schedule

Classes 1-2 — Basics (Chapter 1 of RTP)

Data, Information and Knowledge

- Datum is singular, data is plural (book says “data item” and “data items”)
- Information is data structured and organized to be useful in making a decision or performing some task
- Knowledge implies “understanding” of information
  - Example from book: company analyzes its recruiting data and concludes that recruits from school $X$ tend to have good outcomes only if their GPA’s are at least 3.0. In future, based on this “knowledge”, they screen applicants from school $X$ by their GPA’s, only interviewing those with at least a 3.0 GPA.
  - One common kind of knowledge representation in computers is called “artificial intelligence” (AI). It got a lot of hype in the 1980’s, and then went somewhat out of fashion, but it is still growing gradually. We will not discuss it much, and stick to “information” instead.

Information systems (definition of some basic terms)

- The ways that organizations
  - Store
  - Move
  - Organize
  - Manipulate/process their information
- Components that implement information systems – in other words, Information Technology
  - Hardware – physical tools: computer and network hardware, but also low-tech things like pens and paper
- Software – (changeable) instructions for the hardware
- People
- Procedures – instructions for the people
- Data/databases

- Information systems existed before computers and networks – they just used relatively simple hardware that usually didn’t need software (at least as we know it today). Strictly speaking, this course is about “CBIS” (Computer Based Information Systems). Because of the present ubiquity of these systems, we usually leave the “CB” to be implicitly.

- Impact of computer and network hardware and related software/services (Table 1.1):
  - Can perform numerical computations and other data processing much quicker and cheaper than people
  - Can communicate very quickly and accurately
  - Can store large amounts of information quickly and cheaply; retrieval can often be very rapid
  - Can automate tasks processes that previously required human labor (various degrees possible, of course)
  - Information doesn’t have to be stuck with particular things, locations, or people
  - Can have a downside
    - Small glitches can have much wider impact (minor software bug grounds all aircraft in Japan)
    - Fewer people in the organization understand exactly how information is processed
    - Sometimes malfunctions may go unnoticed (American Airlines yield management story)

- Information architecture is the particular way an organization has arranged its information systems: for example, a particular network of computers running particular software supports the marketing organization, while another network of computers running different software supports the production facilities, etc.

- Information infrastructure consists of the hardware and software that support the information architecture, plus the personnel and services dedicated primarily to maintaining and developing that infrastructure.

- Application and application program are somewhat fuzzy terms, but typically denote computer software and databases supporting a particular task or group of tasks.
  - Example from book: HR uses one application to screen applicants and another to monitor employee turnover
  - A classic business IT problem: applications that don’t communicate with one another (effectively)

Types of information systems:

- Refer to RTP Figure 1.2

- Departmental information systems, or functional area information systems are designed to be operated within a single traditional functional department of an organization such as sales, human resources, or accounting. In the early days of CBIS, these were often the only kind of systems that were practical.

- ERP (Enterprise Resource Planning) systems are a relatively extreme reaction to the problem of poorly integrated functional area systems, offered by vendors such as SAP,
Oracle, and PeopleSoft. They aim to support the entire organization’s needs with essentially a single integrated system. They have enormous potential benefits, but are also notoriously tricky and expensive to configure and install.

- **Transaction Processing Systems (TPS)** gather data about everyday business events in “real time” as they occur. Examples:
  - You buy 3 items at a local store
  - A shipment of coffee beans arrives at a local distribution center
  - A package is unloaded from a FedEx or UPS aircraft
  All of these events are examples of transactions that may be immediately tracked by a TPS. Often, technology like barcodes and scanners makes tracking such transactions, quicker, cheaper, and more detailed.

(End of class 1; begin class 2)

- Some other common terms we will define in more detail later in the course:
  - MIS – “Management Information System”
  - DSS – “Decision Support Systems”
  - EIS – “Executive Information Systems”
- An **InterOrganizational System (IOS)** connects two organizations – for example, it may allow a company to automatically share inventory and backlog data with suppliers or customers.
- **Electronic Commerce of E-Commerce** refers to sales transactions in which at least one side of the transaction (buyer or seller), and perhaps both, is performed by a CBIS without direct human help.

**Role of the Information Systems Department (ISD)**

- Modern computer and network hardware software requires specialized skills and knowledge, at least for firms beyond a certain size.
- This means that the organization needs a sub-organization responsible for IT support: the Information Systems Department (ISD). Names vary from organization to organization; for example, at Rutgers it’s called RUCS.
- In the early days of CBIS, ISD’s “owned” all the information infrastructure because nobody else could understand it.
- As computers became more pervasive and user-friendly, managing IT resources has become a cooperative venture between ISD’s, the departments/functional areas, and “end users” (individuals). Drawing the lines of who is responsible for what can be tricky.
  - At Rutgers, for example, RUCS operates the central network infrastructure, certain key systems like Eden and the outgoing mail server, and works in a consulting/contracting role to support other sub-organizations.
  - Individual or department management of resources will tend to be more responsive and understand user/departmental needs better
  - Central management will tend to have a better understanding of the technology in general, may promote better integration between departments, and can lead to economies of scale
This is a generic management issue that applies to lots of areas besides IT. For example, should each product division have its own product development engineering department, or should multiple divisions share an engineering department?

- In many organizations, the ISD has evolved (or should evolve) into a “business partner” with other departments, and not just a support organization – see below.

**Bulk of Class 2 – Competing in the Digital Economy (RTP Section 1.3 and Chapter 2)**

The “digital economy” described in RTP Chapter 2 has certainly benefited from advances in computing power and information storage. But in the last 10-15 years, the most critical ingredient has been *networking* – the interconnection of multiple computers, and specifically the internet.

A brief history of internet technology (we’ll do a bit more later in the course):

- The key to the present internet is the “TCP/IP” network technology developed in the 1970’s
- TCP/IP was in wide use by research/academic US computer users by the mid 1980’s
- In the late 1980’s, significant research/academic use of TCP/IP began outside the US
- In the early 1990’s, some physicists and computer programmers developed a network-browsable “hypertext” interface called the “world wide web”
- By the mid 1990’s, the WWW drove a massive explosion in internet connectivity and applications; e-mail use “came along for the ride”
- The basic WWW interface was enhanced to gather as well as distribute data
- Technology was developed to link websites to databases
- This technology allowed sales transactions to occur over the WWW
- Physical products are typically delivered by package delivery networks like FedEx and UPS, which experienced symbiotic growth during the 1990’s
  - The idea of a high-performance package delivery network using aircraft was itself pioneered by FedEx in the early 1980’s.
  - IT tools have also been critical to the growth of package delivery networks
- For “digitizable” products like software and music recordings, the product itself could also be *delivered* over the network.

Computer/network-mediated business transactions are called *e-commerce*. Note that the e-commerce is very young; it barely existed 10 years ago.

- Note that e-commerce does not *require* the WWW, although the WWW is a common foundation used to support e-commerce.
- For example, a “B2B” (business to business) e-commerce application in which one firm’s information system automatically communicates parts orders to a supplier’s information system would probably not use the WWW.
- On the other hand, “B2C” (business to consumer) e-commerce applications almost always use WWW interfaces.
Some common terms:

- **The internet**: the global network environment. Literally, it means a bunch of interconnected networks.
- **Intranet**: the network within an organization; typically, it refers to portions of the network not accessible to those outside the organization.
- **Extranet**: one or more interconnected intranets, bridging multiple organizations, but not openly accessible to those outside. For example, a firm might form an extranet with its dealers or critical suppliers, in order to share critical inventory or product lead time information. This information would not be accessible or even detectable just by “googling”.

Some examples of IT issues offering critical opportunities or challenges to companies:

- Obvious examples such as Amazon.com (my own observations; not in this part of RTP)
  - In some cases, a firm’s entire business model is based on IT and the WWW
  - Consider Amazon.com, in their original business of selling books. The key observation is that there are a huge number of different books published, and many books appeal to a “thin”, widely dispersed audience. Thus, any physical store must either limit its selection to better-selling books or carry a huge, slow-moving inventory.
  - The WWW provides an efficient way for customers to browse a huge selection (without mailing out gigantic catalogs)
  - Inventory can be concentrated in relatively few locations, where it turns over relatively quickly
  - Information systems streamline the “picking” and shipping of orders
  - Delivery via efficient package network carriers
  - Amazon.com simply could not exist without modern IT. Another example, of course, is Google.

- **Bringing 7-Eleven out of bankruptcy** (RTP pp.13-14)
  - Old supply system was chaotic:
    - Each store could have more than 80 deliveries per week, each with different items
    - Deliveries could occur during peak shopping hours and disrupt sales.
  - New supply system with stronger IT component:
    - Handheld computers used to place orders
    - Distribution centers consolidate each store’s orders into a single 5 AM delivery the following morning
    - Real time sales and ordering data available to store managers and their superiors
    - Note that this application also involved redesigning the firm’s supply chain. It’s possible to improve supply chains without upgrading information systems, but IT can help a lot.

- **Sarbanes-Oxley at Blue Rhino** (RTP pp. 15-16)
  - Leading supplier of propane canisters for gas grills, sold and collected by independent local distributors.
Since market capitalization exceeds $75 million, the recent Sarbanes-Oxley Act requires both CEO and auditors to certify the firm’s financial system and its controls.

Accounting staff had to plug receivables and payables information from distributors into spreadsheets manually, on a monthly basis.

Resulted in at least one week per month when inventories were not tracked accurately, so the firm had to carry an extra inventory “cushion”

- Classic story at companies that have outgrown desktop computer tools.
- Personal productivity tools like Excel and Word are great, but they are easy to outgrow.
- If many people are repetitively using the same spreadsheet or document, or it is used for routine, cyclic tasks like logging monthly or weekly sales, you have outgrown your desktop tools, and...
- You should invest in a larger scale IT solution constructed with database technology.

HR had to manually communicate information about new hires to the IT department (the process required manual intervention at both departments).

Purchasing required filling out manual forms.

Sarbanes-Oxley accounting controls required all these systems be improved, but the result was also more efficient operations.

- A very similar situation is described in *Crown Media Complies with Sarbanes-Oxley* on RTP p. 51. Again, spreadsheet-based procedures were replaced with an accounting information system using database technology.

- **Kmart and Sears: Ignore Information Systems at Your Peril** (RTP p. 37)

  - While introducing ERP systems may be nightmarish, cobbled-together groups of loosely communicating legacy systems also have serious downsides.
  - Information systems should not just be an afterthought in corporate mergers and acquisitions.
  - Example: Kmart and Sears merged in 2004.

    - **Kmart had**
      - 3 inventory management systems
      - 5 logistics management systems
      - 5 supply chain management systems (the difference between “logistics” and “supply chain” is not entirely clear here)
      - 4 purchasing systems

    - **Sears had**
      - 5 inventory management systems
      - 4 logistics applications
      - 5 supply chain systems
      - 6 merchandise planning systems

  - Each firm’s IT infrastructure was individually a mess; now they have 37 systems to integrate.
  - IT considerations could be a significant obstacle to extracting value from the merger.
While Sears/Kmart struggle with integrating these systems, WalMart and Target can press their advantage through more efficient operations and can get further ahead by adding new technology

Moral: the right IT tools can be a key “competitive advantage”

Consequence of moral: in many organizations, the IT department (ISD) should evolve away from just having a supporting role. Even in a business whose products/services are not directly IT-related, the IT department may need to evolve into a “partner” with (for example) marketing, finance, and/or operations.

Is the moral that all companies should invest aggressively in IT? Not necessarily.

- **Is Dollar General Really Thriving with Minimal IT?** (RTP p. 43): Dollar General has modern IT above the store level, but has minimized its investment in IT at the store level. Individual stores have only cash registers that capture transactions and upload them to the firm’s central IT infrastructure once per night. The firm has had excellent financial performance, although “shrinkage” (theft and lost goods) are bothersome and hard to address due to the lack of in-store information systems.
  - Possible moral: if some aspects of IT are not critical to your business case, they may not merit aggressive investment.

- There are plenty of ERP horror stories to offset the success stories
  - 1999: Hershey reported a $19 million quarterly earnings drop when they brought ERP on line
  - 1999: Whirlpool was unable to ship large numbers of appliances after installing ERP
  - 1996/2001: FoxMeyer (a prescription drug distributor) blames bankruptcy filing on ERP introduction
  - ERP introduction often causes years of dislocation
  - Some firms simply “back out” of ERP introductions after spending millions of dollars.

- There is also no shortage of other IS efforts that have failed or have cost millions more than anticipated
  - 2002: installation of Security Audit and Analysis System (SAAS) at the IRS (RTP pp. 42-43)
  - The FAA introduced the National Airspace System Plan (NASP) to upgrade air traffic control information systems in 1982, but had made little progress by 1991 after huge investments. In 1991, the FAA introduces a more incremental plan that has been much more effective. For example, one element of the incremental plan involved (temporarily) programming new hardware to simulate old hardware which was becoming too difficult and expensive to maintain.

- IT missteps, especially grandiose ones, can be very expensive.

- Sometimes very ambitious, far-reaching IT upgrades that sound great are too difficult to implement or introduce in practice. Gradual upgrade and consolidation of systems may sometimes work better.

One clear message here is that it’s important to think about IT early and try to get it right.

- Changing and upgrading systems can be really painful
• But waiting too long to put in a system is also painful; countless firms are wasting countless employee hours fiddling with spreadsheets when they should have moved to a comprehensive, multi-user database solution long ago.

There’s no magic formula for how firms should approach IT, or whether a particular IT project makes sense. Later in the course, however, we’ll discuss the “SDLC” methodology for evaluating, acquiring, and developing information systems.

I believe some technical understanding of the technology is key to making good decisions. That is why we will start working hands-on with relational database software in the next class.

**Classes 3-4: Basics of Access**

**Class 5: Memory Storage Calculations**

Present-day computer hardware stores all information as combinations of *bits*. A bit can be either 0 or 1, which can be interpreted as “no”/“yes” or “off”/“on”.

Modern processors contain hardware for efficiently manipulating the following collections of bits in one:

- A single bit – 1 means “yes” and 0 means “no”
- 8 bits, also called a “byte” – can hold $2^8 = 256$ possible values. These can represent a single character of text, or a whole number from 0 to 255. If one bit is used to indicate + or –, can hold a whole number from −128 to +127.
- 16 bits, or two bytes. Can hold a single character from a large Asian character set, a whole number between 0 and about 65,000, or (with a sign bit) a whole number between about −32,000 and +32,000.
- 32 bits, or four bytes. Can hold an integer in the range 0 to about 4 billion, or roughly −2 billion to +2 billion. Can also hold a “single precision” floating-point number with the equivalent of about 6 decimal digits of accuracy.
- 64 bits. Can hold a floating-point number with the equivalent of about 15 digits of accuracy, or some really massive whole numbers (in the range of + or − 9 quintillion).

Meaningful manipulation of larger collections of bits is generally a function of software, that is, the way the processor has been programmed (for example, by your creating an application in MS Access).

For the remainder of class 5, refer to the *Memory Storage Calculations* handout.

**Class 6: Data Management (Excerpts from RTP Chapter 4)**

We have seen a little bit now about the tables in which business data are stored, and to how to calculate the amount of storage they might consume.
Chapter 4 of RTP addresses issues of data management. Such issues include:
• How many tables should be in the database an information systems using?
• What data should be in each table?
• How the tables should be connected to one another?
• When an organization has more than one information system (and most do), what information should be in each system?
• How should the systems communicate?

The issues concerning interrelated and overlapping systems resemble in some ways the same questions with tables, but on a larger scale. In this course,
• We will get into a lot of technical detail about how tables should be organized within one database system. It’s a well understood area and the basic concepts are not too difficult.
• As the way entire “systems” should interact, we will take a more descriptive, superficial approach.

• The firm had a very large number of information systems, one for each product line
• Example: if a client informed the firm of a change of address when they updated their automobile policy, that change would not propagate to their life insurance policy
  o This is an example of how redundancy – storing the same data in more than one place – can be a headache for information systems (relate story of safety deposit box)
• Data from all these different systems was hard to integrate, making the firm hard to manage. This problem may have contributed to
  o Poor financial performance relative to competitors
  o Improper sales practices in some units, resulting in a scandal and almost $2 billion in fines and penalties.
• The firm wanted to better integrate business units in response to legislative changes removing barriers between the banking, insurance, and securities industries.
• The company reorganized the data systems serving individual customers with products like life, home, and auto insurance. All these systems now use a common customer database.

MetLife’s problems were not unusual, and many companies have similar problems today. Generically, some common data management problems facing today’s organizations are:
• The volume of data increases as new data is added. Historical data is often kept for a long time, so typically data comes in faster than it is deleted. New technologies mean that gathering new data is easier and faster. So, not only is the total volume of data increasing, but the rate at which it is increasing is also increasing!
  o Should a firm take advantage of every possible opportunity to gather data?
  o For example, any website can gather “clickstream” data: records of exactly how users moved around a website, whether they bought anything or not. This data may be of value, but can pile up quickly.
• Data tend to be scattered throughout the organization. Older organizations tend to have many “legacy” systems that communicated poorly, causing severe problems – the
MetLife case is an example. Thus, it is often desirable to centralize data storage, but by no means always – it may be better to leave departments or working groups “in charge” of the data they use the most. It is costly and risky to replace older “legacy” information subsystems that are working smoothly. Replacing a large number of smaller systems with one larger one can often be very complicated and costly. Sometimes it may be better to created “federated” systems that combine information from constituent systems.

- **Data accuracy** – many organizations have far more errors in their databases than they are aware of. One cause is unnecessary data redundancy (see RTP page 101), but there are other causes too.
- We may also want to use data from outside the organization (either public-domain or purchased).
- It may also be advantageous to share some information with suppliers or vendors (for example, sharing information about inventories can reduce inventory fluctuations and costs throughout a “supply chain”).
- **Data security and quality** are important, but are more easily jeopardized the larger and more complicated an information system becomes.

We can classify the processing tasks information systems perform as follows:

- **Transactional processing** (sometimes called TPS): keeping track of day-to-day events, such as logging orders and shipments, and posting entries to accounting ledgers. In terms of a data table, transaction processing means an ongoing process of adding rows (for example, to reflect a new order), modifying table cells here and there (for example, if a customer changes their telephone number), and perhaps deleting rows.
- **Analytical processing**: means using multiple table rows to obtain “higher-level” information. Entering a row into a table to reflect a newly-received order would be transaction processing; an example of analytical processing would be computing the number of orders and total dollar value of orders for this month, and comparing them to last month.
  - **Analytical processing** can be as simple as sorting, grouping, and summary calculations in an Access query or report. For example, providing all group managers with a summary of their groups’ costs for the month, broken down by cost category. This kind of application can be called “classic” MIS (Management Information Systems).
  - **Analytical processing** can get a lot more sophisticated. For example, **data mining** refers to using sophisticated statistical or related techniques to discover patterns that might not be obvious in classical reports.
  - **Decision support** can involve using the data to help managers make complex decisions, *i.e.* how to route 400 shipments from 20 warehouses to 100 customers.
  - Database systems like Access don’t ordinarily do data mining or decision support by themselves. For such uses, usually need to be connected to other pieces of software.

Sometimes it can be mistake to do transaction processing and analytical processing on the same database, especially if the analytical processing is very time consuming or complex.

- The analytical processing may make the transaction system run slowly.
• Conversely, the transactions may interfere with the analytical processing and make it run to slowly
• If an analytical processing step takes too long, the data it is using may change in the middle of its calculation. “Locking” the data to avoid this can block transaction processing.

It may be better to make a copy or “snapshot” of the database used for the transaction system.
• This is often called a “data warehouse” – see RTP Section 4.4.
• You can do a lot of analysis on the data warehouse without disrupting the transaction system and a lot of transactions without disrupting data analysis
• The data warehouse will not reflect the very latest transactions, but for large-scale aggregate analysis, that may not be a big problem.
• Another reason to create a data warehouse might be to consolidate information from several different transaction systems so it can be analyzed together (see Figure 4.9, RTP page 112).

At this point, we will start getting into the details of data management for the case of a single system. You can think of this material as a very detailed expansion of the subject matter on RTP Sections 4.2 and 4.3.

**Remainder of Class 6: roughly follows GB pages 218-222**
(However, I also introduced the notion of a repeating group.)

**Class 7: Relational Database Design**
See Relational Database Design handout.

**Class 8: Video Store Exercise**
See video store example handout and your own notes.

**Classes 9-10: Multiple Tables in Access**
See personnel records example; in-class work with Access.