1.1 Motivation

1.1.1 Service Economy

Two hundred years ago, if you started a business, it was likely to be a farm. If you could get some land and workers to work that land, you had the basic resources necessary to have an agricultural business. How well you managed those resources would determine how well your farm performed compared to other farms. Most countries’ economies begin as agricultural economies, where farming is the major portion of the economy, and Land and Labor are the most important resources.

One hundred years ago, if you started a business, it was most likely to be a manufacturing business. You would still need land and workers, but you would mainly need a factory. Without the factory, you could not compete against other manufacturers. You would still need some land, but not as much as you would if you were a farm. You would still need some workers, but you would need fewer, better trained workers. The main resource for your manufacturing business would be the factory. Countries whose domestic product comes mainly from manufacturing are known as manufacturing economies, and Capital is the most important resource. Capital might be a factory or machines used to produce items. How well you manage your Land, Labor, and Capital would determine your competitiveness.

If you start a business today, it is most likely to be a service business. You might deliver things (transportation), loan money (banking), connect buyers with sellers, sell things, train, consult, etc. Data is the main resource used to provide a service. You might still need Land, Labor, and Capital, but your most critical resource is data. How well you manage Data will determine how competitive you are.

<table>
<thead>
<tr>
<th>Type of Economy</th>
<th>Agrarian</th>
<th>Industrial</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data is the main resource for today’s developed countries’ economies.

Agriculture-based economy

Industrial economy

Service Economy

GDP Composition by Sector chart
1.1.2 Data Lifespan vs. System Lifespan

1.1.2.1 Example
1.1.2.1.1 Suppose you get a new version of a word processor.
1.1.2.1.2 You “throw out” the old application
1.1.2.1.3 You replace the old application with the new one
1.1.2.1.4 But you keep your documents...
1.1.2.1.5 You open them up in the new application, and it prompts you to do a conversion of the file
1.1.2.1.6 The same thing happens in organizations
1.1.2.1.7 They get a new system
1.1.2.1.8 They “convert” their data for use in the new system (this is called a data migration)
1.1.2.1.9 They throw out the old system

1.1.2.2 Systems have relatively short lifespans

1.1.2.2.1 2-4 years?
1.1.2.2.2 Think of version releases a new system
1.1.2.2.3 Extreme examples
   1.1.2.2.3.1 Cell phone apps
   1.1.2.2.3.2 games

1.1.2.3 Data spans systems

1.1.2.3.1 5 + years
1.1.2.3.2 Space is cheap
1.1.2.3.3 Why throw it out? you might need it.
1.1.2.3.4 Extreme examples
   1.1.2.3.4.1 University registrar
   1.1.2.3.4.2 Law enforcement – criminal records

1.1.2.4 Data grows over time

1.1.2.5 Data has more of a need for management than systems
1.1.3  Data is the basis of decision making

1.1.3.1  No matter how smart we are, with inaccurate data, we make bad decisions.

1.1.3.2  Operations Example

1.1.3.2.1  Suppose we have a certain product that we sell about 100 of per week

1.1.3.2.2  We know that we re-order when the inventory level drops to 20

   1.1.3.2.2.1  We have done extensive analysis to determine that this re-order point will minimize our holding cost and also keep lost sales to a minimum

1.1.3.2.3  Suppose our system says there are 40 in inventory, when there are really 5?

1.1.3.2.4  Bad decision (don’t order more) based on bad data

1.1.3.3  Marketing Example

1.1.3.3.1  Suppose we have introduced a new product

1.1.3.3.2  The system says we have sold 2000 units over the last week

1.1.3.3.3  We need to be selling 4000 units per week to make a profit

   1.1.3.3.3.1  We have done extensive analysis to determine that 4000 units is our break-even point.

1.1.3.3.4  We discontinue the product

1.1.3.3.5  In actuality, we sold 10,000 units last week

   1.1.3.3.5.1  Maybe there is a delay in reporting sales

   1.1.3.3.5.2  Maybe there was a typo – somebody entered a negative number for units sold

   1.1.3.3.5.3  Maybe the data was scrambled as it moved from one system to another

   1.1.3.3.5.4  Whatever...

1.1.3.3.6  Bad decision (discontinue product) based on bad data

1.1.4  Information Overload

1.1.4.1  Data is gathered at unprecedented levels

   1.1.4.1.1  Every scanner at Walmart stores every item that is scanned

1.1.4.2  As capabilities increase, more detail is stored
1.1.4.2.1 We used to just store the total of each sale (shopping cart)

1.1.4.2.2 Now we store

1.1.4.2.2.1 Every item in the cart

1.1.4.2.2.2 Coupons/discounts

1.1.4.2.2.3 Who bought it

1.1.4.3 Management of data is essential

1.1.5 Flattening of Organization Pyramid

People in early organizations were arranged in hierarchies. A hierarchical organization chart looks like this the following figure. Notice that Workers report to Managers, who report to Directors, who report to Vice Presidents, who report to the President.

The hierarchy represented the reporting structure (who reports to who), but also the flow of information. Individual workers knew only their position, and were the best equipped to make their own decisions. Managers knew information about all workers under them, but not about workers under other managers. Managers were the best informed to make decisions about all the workers under them. Directors knew information about managers under them, but not about managers under other directors. Furthermore, Directors did not know details about workers under them. Directors knew generalities about each Managers’ workers. For example, Directors don’t know how many widgets each worker produced, only how many total widgets were produced by a manager’s workers.

To some extent the hierarchy defined the information that a person in the organization had. Because information flow was limited by technology, the management span-of-control was relatively small. Furthermore, workers were very limited in their decision making capability, due to their limited access to information. This lead to some interesting organization dynamics and sayings such as “knowledge is power”. What if a Manager had all the information of her Director? Could she take her director’s job?
1.1.5.1 **The hierarchy has fewer levels**

Today’s organizations tend to have fewer levels. A more modern version of the organization above might be:

Notice that there are fewer layers between the top and bottom of the organization. Furthermore, there are fewer “middle-management”. The Director layer is completely gone, resulting in 4 fewer directors, and the Manager layer now has only 4 Managers. This change is at least partially due to easier
information flow enabled by modern technology. It’s quite common in today’s organizations for all levels of the hierarchy to have access to all information. Because of easily accessible information, workers are more able to make the types of decisions that only managers could make before. Workers can be more effective independently, which means that they need less guidance from their managers, which means that managers can manage more workers; the management span-of-control has increased.

For new entrants in the work force, this means that they are likely to be decision-makers from their first day. In the past, entry-level workers would “do what they were told”, implementing decisions made by their superiors. Today, entry-level workers will decide on their own what to do. And why not? They have access to all the information of their superiors. How well entry-level workers perform compared to their peers will depend on how well they manage the information accessible to everyone.

1.1.6 Database is a job-winning skill

1.1.6.1 Every discipline uses data for decision making
1.1.6.2 Every organization uses data for decision making
1.1.6.3 The next step past spreadsheets
1.1.6.4 Given two equivalent candidates, the one with database experience has the edge

1.1.7 History Lesson – Programs and Files

1.1.7.1

1.1.7.2 Problems with File-based legacy systems
1.1.7.3 How do RDBMS address these problems
1.1.7.4 What about other storage mechanisms?

1.1.7.4.1 File systems
1.1.7.4.2 Hierarchical DBMSs
1.1.7.4.3 Network DBMSs
1.1.7.4.4 OODBMSs

1.1.8 What are the advantages of RDBMSs?

1.1.9 What are the disadvantages?
1.2 Data & Information

1.2.1 Input, Processing, Output

1.2.2 Data
  1.2.2.1 Definition
  1.2.2.2 Attributes
  1.2.2.3 Examples

1.2.3 Information
  1.2.3.1 Definition
  1.2.3.2 Attributes
  1.2.3.3 Examples

1.2.4 Operations that might convert Data to Information
  1.2.4.1 Filtering
  1.2.4.2 Sorting
  1.2.4.3 Formatting
  1.2.4.4 Changing representation
    1.2.4.4.1 Charts, graphs, images, etc.
  1.2.4.5 Aggregation
  1.2.4.6 Changing Time of Presentation
  1.2.4.7 Changing Context of Presentation
  1.2.4.8 Integrating related data
  1.2.4.9