* Definitions （and all other types of business rules）are gathered from the same sources as all requirements for information systems . Thus, systems and data analysts should be looking for data objects and their definitions as these sources or information systems requirement are studied .
* Definitions will usually be accompanied by diagrams, such as entity-relationship diagrams .The definition does not need to repeat what is shown on the diagram but rather supplement the diagram.
* Definition will be stated in the singular and explain what the data element is , not what it is not .A definition will use commonly understood terms and abbreviations and stand alone in its meaning and not embed other definitions within it .It should be concise and concentrate on the essential meaning of a data object as
  + - Subtleties
    - Special or exceptional conditions
    - Examples
    - Where , when ,and how the data are created or calculated in the organization
    - Whether the data are static or changed over time
    - Whether the data are singular or plural in their atomic form
    - Who determines the value for the data
    - Who owns the data (who controls the definition and usage )
    - Whether the data are optional or whether empty (what we will call null) values are allowed
    - Whether the data can be broken down into more atomic parts or are often combined with other data into some composite or aggregate form

If not included in a data definition, these characteristics need to be documented elsewhere,where other metadata are stored.

* A data object should not be added to a data model,such as an entity-relationship diagram,until after it has been carefully defined(and named) and there is agreement on this definition. But expect the definition of the data to change once you place the object on the diagram because the process of developing a data model tests your understanding of the meaning of data.(In other words, modeling data is an iterative process.)

There is an unattributed phrase in data modeling that highlights the importance of good data definitions:”The person who controls the meaning of data controls the data.”it might seem that obtaining concurrence in an organization on the definitions to be used for the various terms and facts should be relatively easy.however,this is usually far from the case.in fact,it is likely to be one of the most difficult challenges you will face in data modeling or,for that matter,in any other endeavor.it is not unusual for an organization to have multiple definitions (perhaps a dozen or more )for common terms such as customer or order.

To illustrate the problems inherent in developing definitions,consider a data object of student found in a typical university.a sample definition for student is “a person who has been admitted to the school and who has registered for at least one course during the past year .”this definition is certain to be challenged,because it is probably too narrow.a person who is a student typically proceeds through several stages in relationship with the school,such as the following:

1. prospect-some formal contact,indicating an interest in the school
2. applicant-applies for admission
3. admitted applicant-admitted to the school and perhaps to a degree program
4. matriculated student-registers for at least one course
5. continuing student-registers for courses on an ongoing basis(no substantial gaps)
6. former student-fails to register for courses during some stipulated period(now many reapply)
7. graduate-satisfactorily completes some degree program(now many apply for another program)

imagine the difficulty of obtaining consensus on a single definition in this situation! It

1. ***Use multiple definitions to cover the various situations.***this is likely to be highly confusing if there is only one entity type ,so this approach is not recommended (multiple definitions are not good definitions)
2. ***use a very general definition that will cover most situations.***this approach may necessitate adding additional data about students to record a given student’s actual status.
3. ***consider using multiple,related data objects for students.***for example,we could create a general entity type for student and then other specific entity types for kinds of students with unique characteristics.we describe the conditions that suggest this approach in chapter3.

**MODELING ENTITIES AND ATTRIBUTES**

The basic constructs of the E-R model are entities,relationships,and attributes.as shown in figure 2-2,the model allows numerous variations for each of these constructs. The richness of the E-R model allows designers to model real-world situations accurately and expressively,which helps account for the popularity of the model.

**Entities**

An entity is a person,a place,an object,an event,or a concept in the user environment about which the organization wishes to maintain data.thus,an entity has a noun name.some examples of each of these kinds of entities follow:

Person: EMPLOYEE,STUDENT,PATIENT

Place: STORE,WAREHOUSE,STATE

Object: MACHINE,BUILDING,AUTOMOBILE

Event； SALE，REGISTRATION，RENEWAL

Concept: ACCOUNT,COURSE,WORK,CENTER

ENTITY TYPE VERSUS ENTITY INSTANCE There is an important distinction between entity types and entity instances. an entity type is a collection of entities that share common properties or characteristics.each entity type in an E-R model is given a name.because the name represents a collection(or set) of items,it is always singular.we use capital letters for names of entity type(s).in an E-R diagram,the entity name is placed inside the box representing the entity type(see figure2-1).

An entity instance is a single occurrence of an entity type.figure 2-3 illustrates the distinction between an entity type and two of its instances.an entity type is described just once(using metadata)in a database,whereas many instances of that entity type may