

Process Modeling: Context Diagrams and Data Flow Diagrams (DFDs)

Introduction


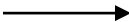
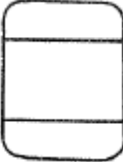
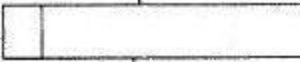
Figuring out the business processes for complex systems can be complicated. For example, if the goal is to streamline an existing supply chain process, your investigation will cross multiple business units, perhaps starting with an on-line order, a retail store pick-up, or a telephone order. How does the current process work to replace items in inventory storage and on retail shelves? Computer systems make these processes faster, but often increase complexity. As a business analyst, programmer analyst, or IT auditor you are likely to need a graphical technique to help investigate and document current processes and work with a team to determine where problems occur and what the

best solution is. Graphical process models are a common part of Root Cause analysis to determine exactly where problems occur. Graphics help teams communicate what software needs to be created or fixed based on a view of what data must be processed to meet and fulfill system requirements.

Description

A **data flow diagram (DFD)** is a drawing that shows how a system's environmental entities, processes, and data are interconnected. Using four simple symbols users can show developers their current system processes and what they would like to change. The only four symbols are shown in Table 1 below.

Table 1. Data Flow Diagram Symbols

	<p>A square is an Environmental Entity (EE)—a source of, or destination for, data outside the system.</p>
	<p>An arrow is a data flow. Each data flow must have a unique identifier.</p>
	<p>A process bubble is a process that changes data. Process bubbles should be numbered at the top. The Process Label should be in a verb – object format. The bottom section identifies an actor or system component.</p>
	<p>An open rectangle is a data store.</p>

The Context Diagram

DFDs exist in a hierarchy, beginning with the simplest, highest, system level and ending with the lower level diagrams with detailed processes.

The DFD at the system level illustrates the *context*, that is, the circumstances of its environment and is called the **context diagram**. This diagram contains only a single,

unnumbered process bubble in the shape of a circle to represent the entire system.

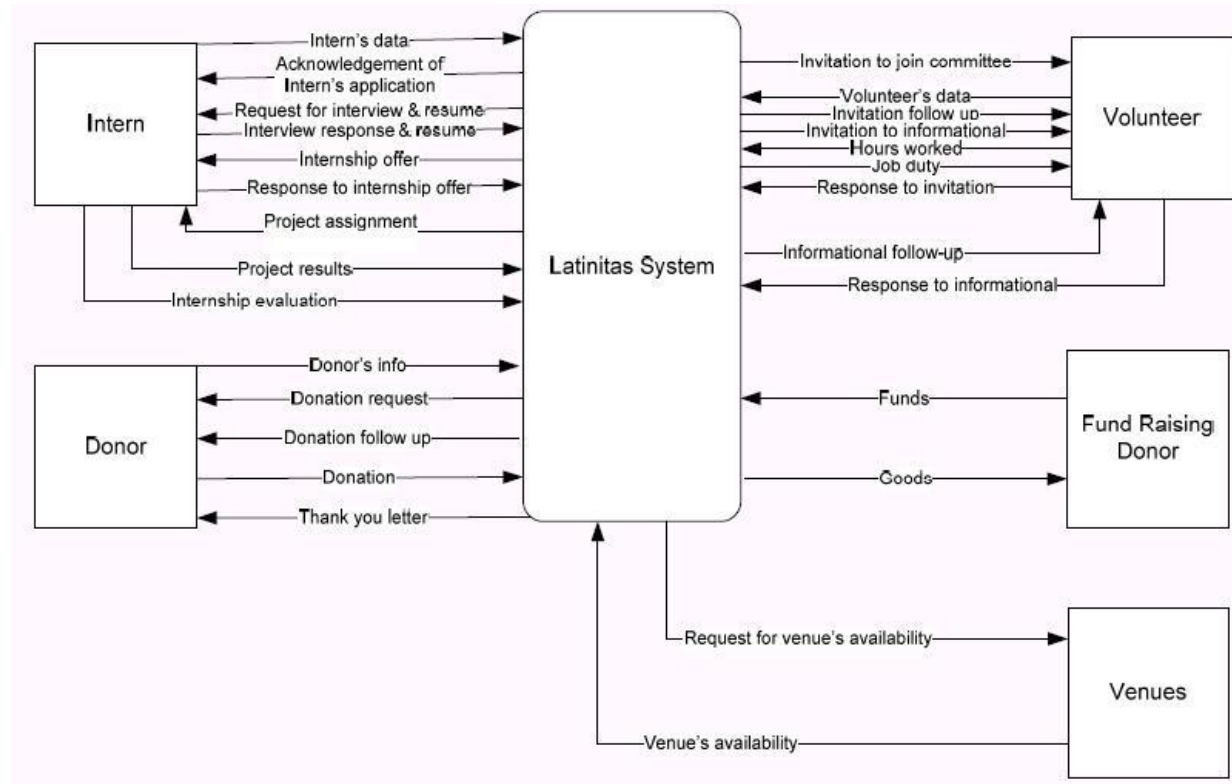
Figure 1 shows the context diagram for the initial investigation of the *existing* Latinitas “system” with a single process bubble in the middle. Its five environmental entities are represented by squares surrounding the single system process bubble. The system is connected to its environmental entities by arrows that represent data flows. Notice how many data flows



surfaced in the team’s discussions with Angie about her work with Latinitas donors, volunteers, and interns. When Angie posted her request for an MIS 374 “Volunteer Database” she had not

thought through her real needs. White-boarding with her MIS 374 development team was a learning experience for Angie as well as the team.

Figure 1. Latinitas Current System - Context Diagram



The Figure 0 Diagram

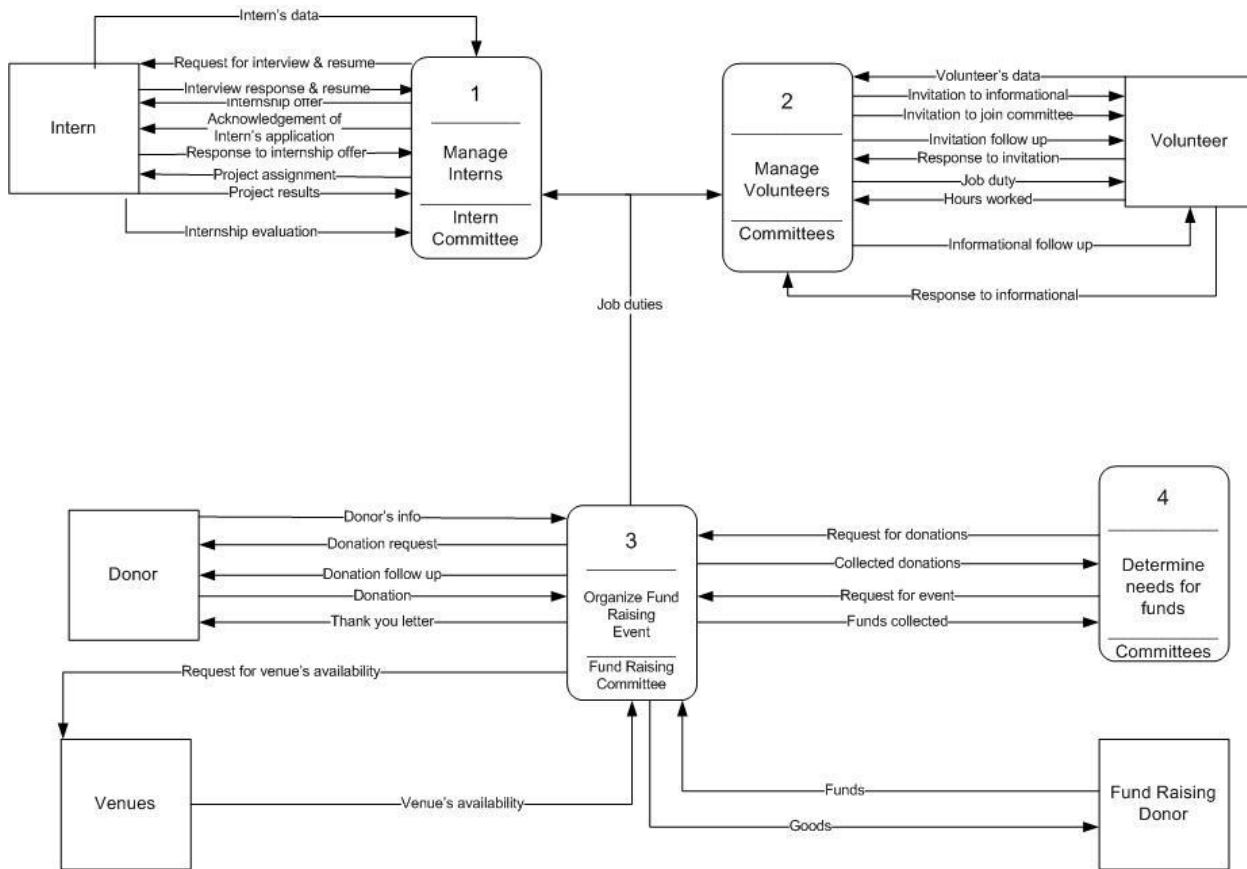
A DFD diagram on a lower level “explodes” a process on the level immediately above, providing more detail. The DFD on the second level from the top in the hierarchy is the **Figure 0 diagram**. Each process of the Figure 0 diagram is numbered from left to right and then down. In order to prevent a DFD from becoming cluttered, the general rule is to keep the number

of processes to seven or less. This rule applies not only to the Figure 0 diagram but to DFDs on the lower levels as well.

Figure 2 shows the Figure 0 diagram of the Latinitas system. The system contains four main processes: Manage Interns, Manage Volunteers, Organize Fund Raising Events, and Determine Needs for Funds.



Figure 2. Latinitas System – Figure 0 Diagram



Note that this Figure 0 diagram includes the same five environmental entities from the context diagram and that the data flows connecting these entities with the processes all have the same labels as in the context diagram. For example the Fund Raising Donor at the bottom right of Figure 0 above has exactly two data flows with the same names as the two data flows for the Fund Raising Donor in the Context Diagram.

Figure n Diagrams

As you continue the process of documenting data flows, DFDs with such names as Figure 1 diagram, Figure 2 diagram, and so on are created. The Figure 1 diagram documents the major processes of Process 1 of the Figure 0 diagram; the Figure 2 diagram documents Process 2, and so on. We refer to the DFDs on this and additional levels as **Figure n diagrams**. Figure 3 below illustrates a Level 1 diagram. It documents the major Process 3 of the Latinitas

system—Organize Fund Raising Event. This exploded DFD consists of five detailed processes numbered 3.1, 3.2, etc. Process 3.1 is the initial process of obtaining and processing donations needed to hold an event. Process 3.1 is exploded in Figure 4. A Level 2 (Figure 3.1) DFD for Obtaining Donations is on page 5. This is a Level 2 diagram. Notice that the process numbers for the exploded detail process are 3.1.1, 3.1.2, 3.1.3, and 3.1.4.



Figure 3. Latinitas - Organizing Fund Raising Events Process

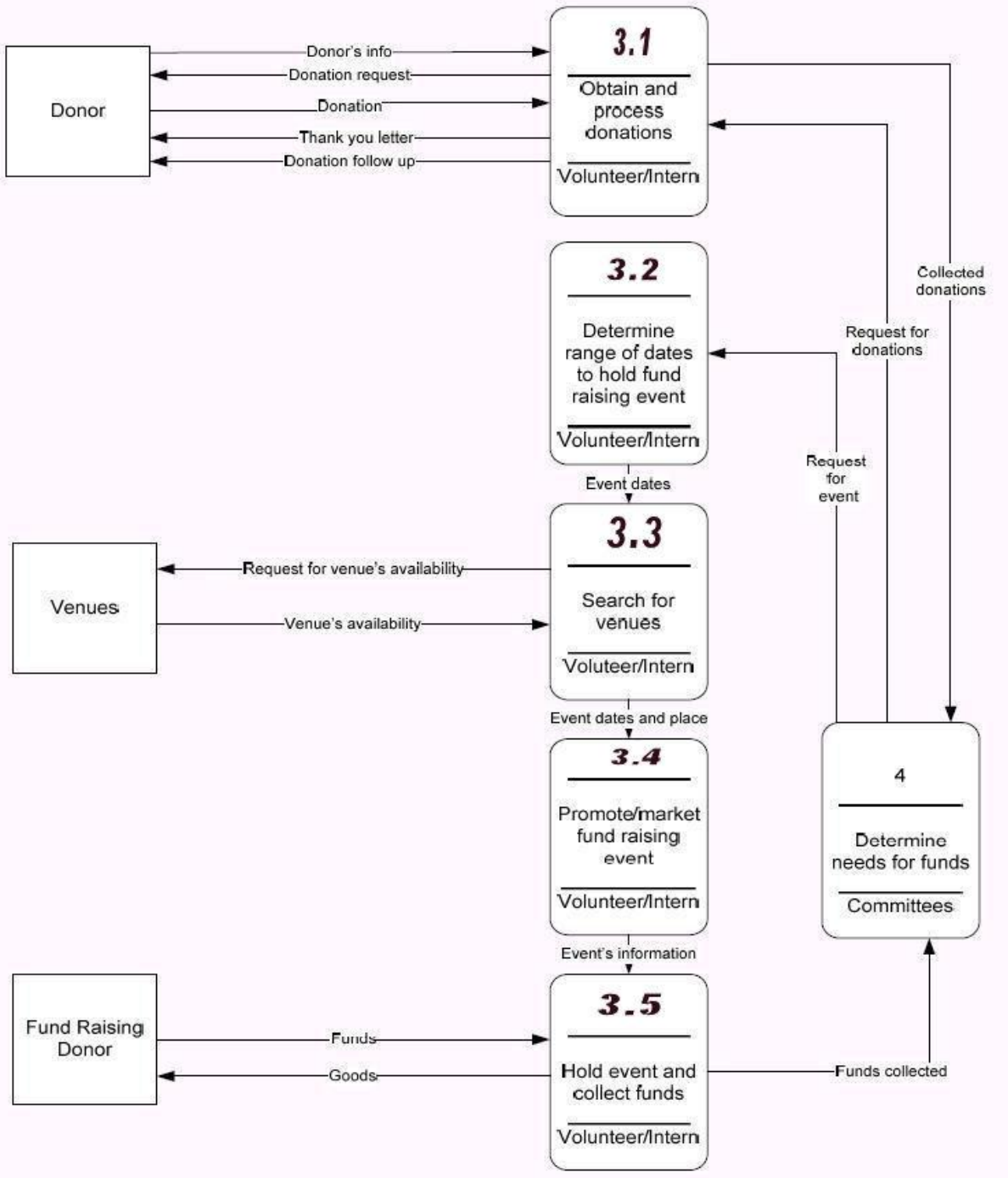
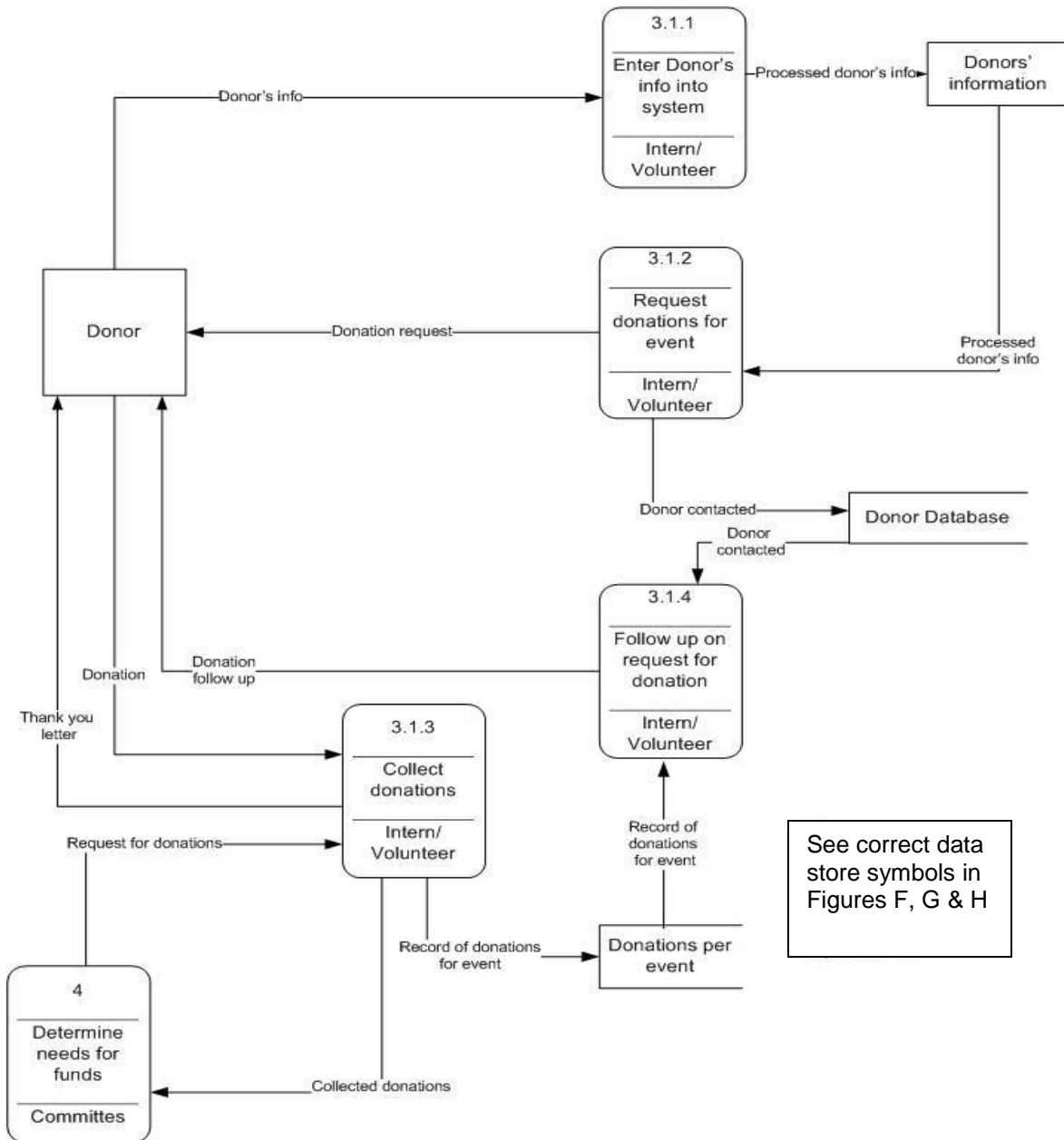




Figure 4. Latinitas – Obtaining Donations Process



Notice that a 4th symbol – a data store has been introduced at this level of process documentation for Latinitas. “Donor Information” is a manila folder that Angie keeps; the “Donor Database” and “Donations Per Event” are spreadsheets. These three data stores should have identifiers, such as D1, D2 and D3, depending on how many file folders and spreadsheets Angie and others create in their work. (Figures 6, 7, and 8 have correctly identified data stores.)

The same top-down decomposition of the system processes can be continued for more detailed levels. However, it is not necessary to document the same level of detail for all processes. Some processes can be adequately documented at a higher level in the hierarchy. The top-down process of documenting a system with DFDs continues until you reach a level of detail where the focus is on individual data flows that are important to determining functional requirements for the scope of a project.

Creating DFDs

Our example figures for Latinitas are the result of considerable time spent by the team with Angie, their client. A good way to start is to create the simple context diagram with a circle in the middle. Then ask your client about what data the system must produce and what data is required for the system's functionality. If you do your work on a white-board, photograph the board at the end of your discussion. If you think this is fairly accurate, then create a version in Visio that will allow you to add Environmental Entities and data flows as you talk with stakeholders to learn more details and create lower level DFDs.

Few clients will be able to provide all your answers, so you will need to talk with key stakeholders that know the existing system to analyze the processes for an existing system. Individuals are likely to be able to tell you only their processes, so you start with a "Figure 0" DFD that you label with some meaningful name. You might end up with a dozen of these "Figure 0" diagrams. At this point your team wrestles with all your messy diagrams and tries to come up with a numbering system that makes sense for your client. This is what the Latinitas team did to create their Figure 0 diagram (see Figure 2 on page 3 of this document.)

DFD Symbols and Rules

As stated on page 1, DFDs consist of four basic symbols that represent (1) environmental entities, (2) data flows, (3) processes, and (4) data stores.

Environmental Entity (EE) Symbols. Systems interface with persons, organizations, locations, and other systems. Each of these entities in the environment is represented by a square and includes the name of the entity. Sometimes

drawings become cluttered if data flow arrows are needed from several different processes to the same EE. In this case, consider repeating an EE square; use the naming convention of EE1/1 and EE1/2 to indicate that EE1 has a 1st and 2nd instance. (The Patient entity is repeated and identified this way in Figure G.)

Process Symbols. A process is a transformation of data. Data flows into a process, is transformed, and then flows out. In order for something to be a process, the output must be different from the input in some way. Process symbols are labeled with a verb and object, such as Verify User Name and Edit Date. The single process bubble in a context diagram is an exception: it is labeled with the system name.

Data Flow Symbols. A good way to think of a data flow is "data on the move." The data moves from an environmental entity to a process, from one process to another, from a data store to a process, and so on. In most cases the data flow is a single line going in one direction; two-headed (or forked) arrows can be used to show the same data flows from one process, data store, or entity to more than one destination. (For example, in Figure 6 the Monthly Newsletter flows from Process 2 to both the Website Database and to Referring Physicians.) Each data flow in a DFD is labeled with a *unique* name in the same diagram. However, when the same flow appears on more than one level the same name is used for each appearance.

Data Store Symbols. Whereas data flows represent data on the move, data stores represent data that is maintained in fixed locations. Think of a data store as "data at rest." Examples of data stores are master files, such as inventory and personnel data that are kept current and history files that are held in archival storage.

Data stores are depicted with open-ended rectangles and have a unique label, such as D1, D2, etc. or S1, S2, S3, etc. as illustrated in Figures 6, 7, and 8.



The Story-Telling Power of DFDs

DFDs tell stories. They capture real-world business processes in a way that enables you to communicate with other IT professionals and stakeholders. Therefore, another trained person (i.e.: a 374 classmate, the TAs, and the professors) should be able to read your DFDs without ever reading the supporting case materials. Your DFDs should be self-

explanatory and thoroughly explain the system's processes. Read through the three levels of process models in Figures 5 through 8. These were created by a past MIS 374 team for Central Texas Pediatric Orthopedics (CTPO). Following the story should be easier for this web development project than the more complicated Client Relationship Management (CRM) System for Latinitas.

Figure 5.

CTPO: Proposed Content Management System: Context Diagram

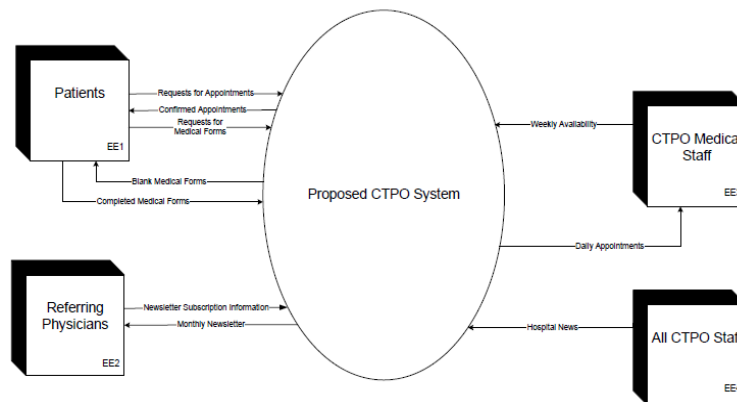
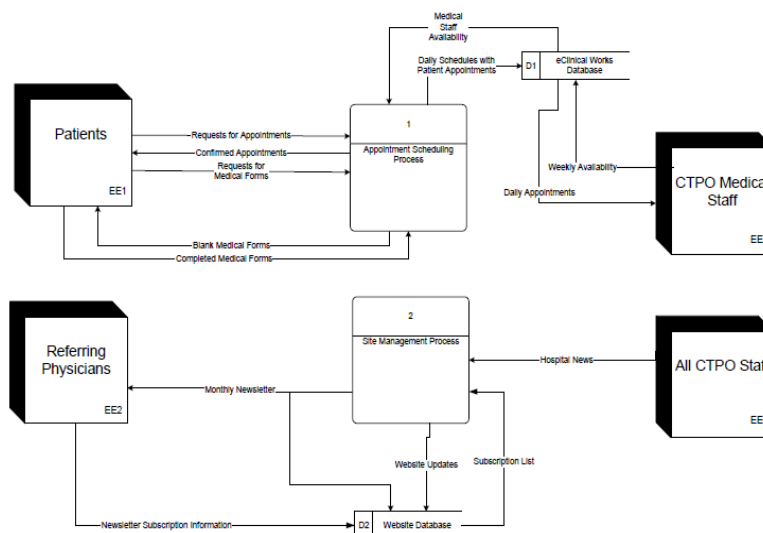


Figure 6.

CTPO: Proposed Content Management System: Figure 0

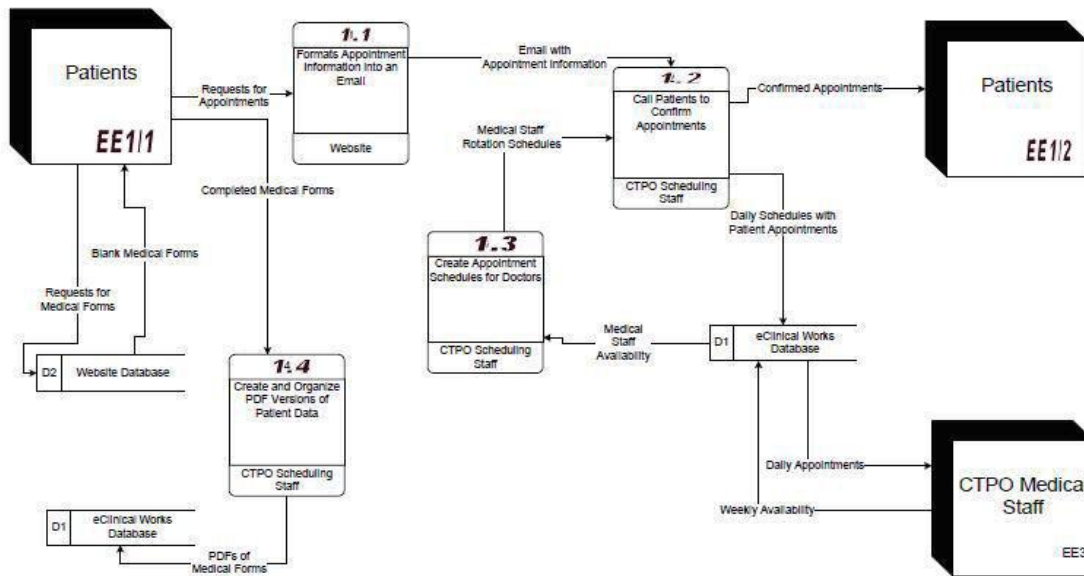


Note that these process bubbles on the Figure 0 do not include actors. This is correct, since each process is exploded to show that the detail steps are performed by different people or system steps.



Figure 7.

CTPO: Proposed Content Management System: Figure 1



FAQs

Q1: Where can I find the Visio stencils used in the examples in Figures 6 through 8? **Answer:** The Visio Stencil Instructions and the Visio Stencil are available on the Resources page of the class web site in the Process Modeling block. Read the "Instructions" file, and download the .zip file. Open the stencil within Visio and the stencils (environmental entities, processes, data stores, etc.) will become available. For data flows, use right-angle connectors.

Q2: How do I even begin to make my DFDs? I'm totally stuck. **Answer:** Try listing major processes and environmental entities. Ideas for EEs should come from your Stakeholders Table. Ideas for the major processes should come from your Root Cause Analysis. From this point try the approach listed for the next FAQ – put the EEs on the edge of the paper and the process bubbles inside for a rough start on a Figure 0

diagram. Or start with lower level processes with the same approach. Make a list of all the data flows that go to and from the environmental entities in your rough level 1 diagrams, so you check that they all appear on the Figure 0 diagram and Context diagram.

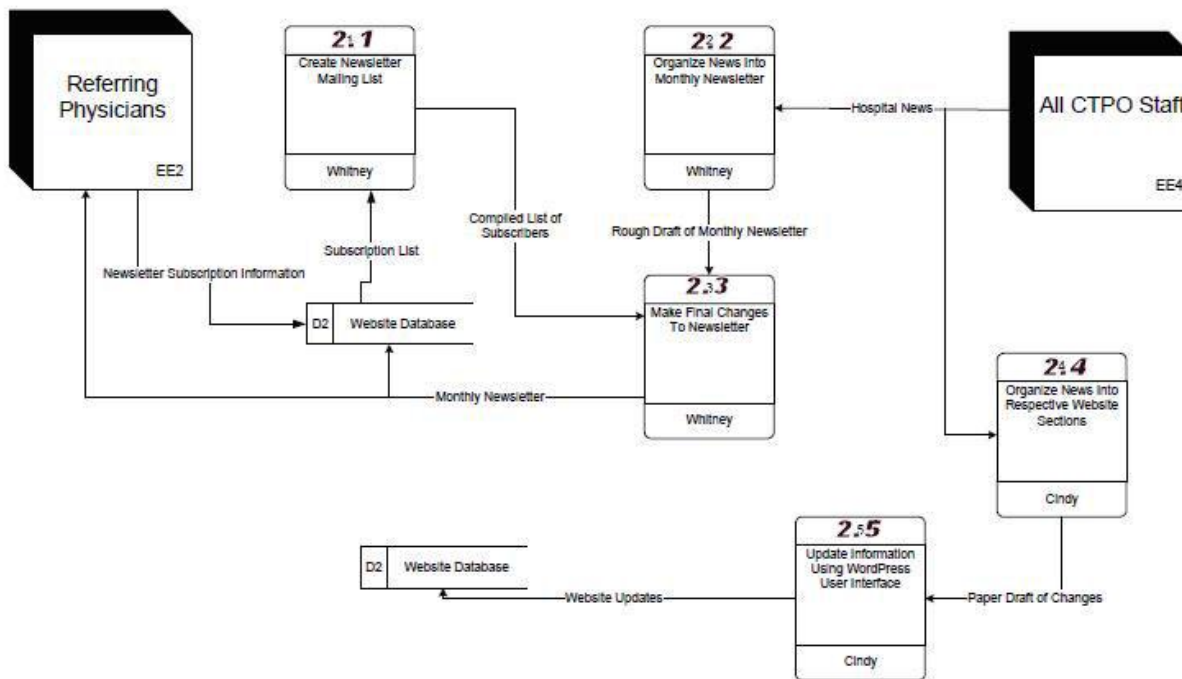
Q3: How do I make my DFDs pretty? All the lines are crisscrossed. **Answer:** Try to organize your environmental entities (EEs) around the outside edges of the paper. Then, put your process bubbles inside. If any process uses multiple EEs, try to put the EEs close to each other. Also try to use right-angle connectors; they make the diagrams more legible, as well as, help you group in-bound and out-bound data flows together. Check the Visio help file (F1 key) for more information on using right-angle connectors.

Some analysts find it easier to work on paper before using Visio, but some analysts do it the other way around. If you make lists, then you may be able to jump straight into Visio and start



Figure 8.

CTPO: Proposed Content Management System: Figure 2



diagramming, at least for the Figure 0 and Context Diagrams—after you have completed your Figures 1, 2, and 3.

For your client project you may want to white board processes and then take photos to use for Visio graphics, but for your Group Project 1 case study (and any time you have text as the basis for your graphics), your choice is hand-drawn or Visio as your first step.

Q4: What are some general hints when making DFDs? **Answer:** Label your environmental entities (ex: EE1 and EE2). If you use the same environmental entity multiple times, make sure each has the same EE number. Also, make sure environmental entities don't "talk" with each other. That is, don't draw data flows between environmental entities – instead, these data flows must be connected through some process bubble in the System.

Remember that every process must transform data. Printing doesn't transform data. A data

flow arrow, not a process bubble, shows sending or delivering data.

Use the Gane-Sarson process symbols (provided in the Visio Stencil for MIS 374). This helps to clarify who/what performs the processes. This makes it easier for the TAs to understand what you are trying to communicate, in case of discrepancies.

NOTE: Gane-Sarson process symbols are the process bubbles above with three parts – process ID number on top, verb – object process label in the middle, and who/what performs the process on the bottom.



High Quality Delivery Tips

Data Flow Diagrams:

- Include a title that states what system process is being modeled and whether it is an existing system or a proposed system
- Number each DFD Figure.
- Environmental Entities (EE):
 - Labeled with EE number
 - Same EE number and name label on each diagram
 - If used twice on same diagram, label EE 1/1 and EE 1/2
 - No data flows between EEs
- Data Flow:
 - Arrow pointing in direction of data flow
 - Data flows must match exactly between higher-level and lower-level DFDs
 - No two data flows should have the same name
- Process Bubble:
 - Processes must transform data
 - Must be numbered
 - Numbering must be consistent across all levels of DFDs
 - Must have verb-object format
 - Actors should be included, but not necessary for Figure 0 DFD
 - No more than 7 processes per DFD
- Data Store:
 - Open rectangle
 - Must be labeled with S number or D number
 - Consistent across all diagrams and levels
 - Do not communicate with each other or EEs directly

Context Diagram:

- Include a title that states what system is being modeled and whether it is an existing system or a proposed system.
- One single, unnumbered process bubble in shape of circle- labeled with system name or verb-object phrase
- Show only System, Environmental Entities, and Data Flows
- Data flows and EEs in Context Diagram must match *exactly* to those in the Figure 0 DFD

Examples on [Resources Page](#)

- [DFD Summary and Rules](#)
- [DFD Examples w/ Tips](#)
- [PrelimDonorProcess](#) (Visio vsd file) for Latinitas