

Data Flow Diagram

Originally prepared for the CDFW-BDB¹ Field Data Collection Workgroup, 11/12/2007
Updated by Douglas Burch, 12/5/2014

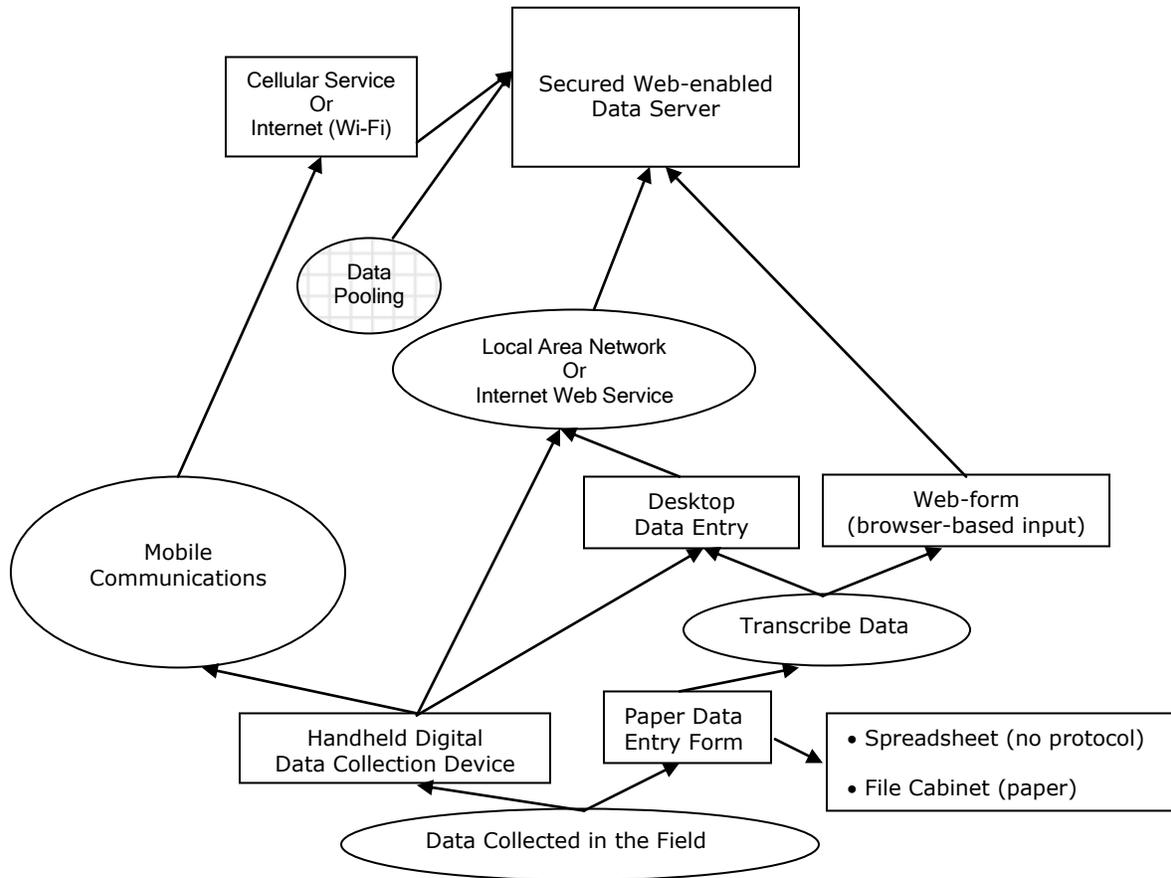


Figure 1. A flow chart of some of the more common options implemented when considering the one directional flow of raw field data to a secure environment where it can be stored and managed most efficiently.

¹ BDB is the acronym for Biogeographic Data Branch which is a branch of the California Department of Fish and Wildlife's (CDFW) Data and Technology Division.

There are a number of possible ways for field data to flow from its source to where it is ultimately resides (Figure 1). The optimal path will depend on a variety of factors. Below are a few of the considerations:

1. Requirements of the project
2. Hardware and software available to the project (budget dependent)
3. Technical skills of the project staff
4. Time allowed by project for technical skills to be learned
5. Level of support from Information Technology (IT) staff
6. The degree to which project managers are comfortable (willing to use) technological approaches to managing data

Methods of Data Flow

1. **Web-based Data Entry System** – Input is entered directly to a web-based application. This is an efficient data flow path, because there is minimal delay before the entered data can be available almost anywhere. Therefore, managing applications on web-enabled data servers is much more efficient than with distributed database applications which require merging or pooling of the data before it can be made available all in one place. The drawback is that web-based data entry applications can be cumbersome and slow the input of data making this a poor option when large amounts of field data must be entered.
2. **Traditional Field Data Entry System** – Field crew enters data onto paper forms in the field and later transcribes that data to a desktop data entry interface connected to a back-end database. If there is a desire for more than one desktop data entry interface per project then the back-end database data might require one of the following options to be met:
 - a. It has a networked connection to the frontend data entry interface or application forms.
 - b. Data compiled from local databases on separate disconnected workstations must have their data pooled or merged before a common set of project data is available for inclusion into a central back-end database.
 - c. Additionally, without a direct connection to the back-end central database pooled or merged data will require a database data manager

available for receiving and appending to the central database where it can be made available by way of a web-enabled server.

3. **Handheld Digital Data Entry System** – Entering data directly into a handheld device while collecting data in the field eliminates the need for transcribing data a second time, however many data project managers will still want to use paper forms as a fallback system or as an adjunct to a handheld digital data entry system when a project does not have sufficient numbers of handheld devices.
 - a. Handheld digital devices with mobile configurations can upload to a central data server if the infrastructure is in place, however having access to these IT resources might be a challenge for some projects, and some data may still need to be entered in the traditional way.
 - b. Handheld digital data collection devices can be connected to and synchronized with desktop applications on local area workstations.
 - c. Entering data in to a handheld device – only one time – while still in the field collecting the data eliminates the redundancy involved when data is reentered or transcribed from paper datasheets. In evaluating issues related to data quality, detected levels of error associated with transcription have been shown to be as much as 44% of the overall data error and suggest that that this is just as relevant as data error that is the result of the methodology used to enter the data (Johnson et. al. 2009); based on this information it stands to reason that by eliminating the second transcription of raw field data, i.e., from paper to electronic format, one would also reduce this potential risk of human error from happening more than once.
4. **Local Area Network / Internet Web Services** – This method allows desktop data entry applications to interface with a web-enabled data server. It requires varying levels of infrastructure and support, but provides multiple options as far as getting large amounts of data from handheld devices or desktop computer applications to a web-enabled data server.

CITATION

Johnson, C.L., G. Temple, T. Pearsons, T. Webster. 2009. An Evaluation of Data Entry Error and Proofing Methods for Fisheries Data. Transactions of the American Fisheries Society 138:593-601