



GUIDELINES FOR DATABASES AS PUBLIC RECORDS

PURPOSE	1
OVERVIEW	1
POLICY GUIDELINES	2
<i>OFFICIAL REQUEST</i>	2
<i>EXEMPT RECORDS</i>	2
<i>REQUESTS FOR SPECIFIC RECORDS</i>	3
<i>REQUEST FOR ENTIRE DATABASES OR FOR PORTIONS OF DATABASES</i>	3
RECOMMENDATIONS	4
DEFINITIONS	4

PURPOSE

The purpose of these guidelines is to assist State Agencies when responding to public records requests when the requested public records are contained in an electronic database. These guidelines are general in nature, they are not intended to address all the various issues that can arise in responding to such requests, and they are based on Ohio's public records laws as of December 2001. Specific questions on an agency's obligations and responsibilities in responding to public records requests should be addressed to in-house counsel or to the Office of the Attorney General.

OVERVIEW

In today's technological environment, information is routinely maintained and stored in electronic databases. Ohio's Public Records Act, located in Section 149 of the Ohio Revised Code, requires that public entities make records¹ contained within electronic databases available to the public upon request.

Generally, requestors have specific needs and require only the generated output of databases to obtain the records or information being sought. These outputs are routinely filtered or sorted through report queries or standard reporting processes to present information in a meaningful manner to the user.

The Public Records Act does not require an agency to create new records by searching and retrieving information from pre-existing records or databases. However, if there is an existing query, filter, or sort then the record already exists for the purposes of the Public Records Act.

Because of this, a great challenge exists when public records requests are for an entire database or portions of a database. A database can consist of many inter-dependent components. Such components can include, but are not limited to, the following: software, hardware, program logic, data tables, security tables, access controls, data and table links, mathematical and other logical computations, etc. In some cases, records are stored in multiple databases and on multiple operating platforms, which further complicates the issue. In essence, raw data or records are not always useful without the proper components of the database being linked together or made available.

GUIDELINES FOR DATABASES AS PUBLIC RECORDS

OHIO ELECTRONIC RECORDS COMMITTEE

<http://www.ohiohistory.org/ohiojunction/erc/>

The challenge is to provide records as legally required by taking reasonable measures to extract the records requested without compromising system security or providing proprietary information which may be in violation of licensing agreements.

POLICY GUIDELINES

OFFICIAL REQUEST

A contact person at your agency should help coordinate the public records request to ensure that the requestor is satisfied with the information obtained. Although not required, whenever possible, public records requests should be obtained in writing to ensure there is no misunderstanding of the requestor's needs. If the requestor will not complete a written request, then the contact person should record the request to ensure accuracy.

The request should be as specific as possible and may require a thorough interview process to determine what information is being requested. The contact person should ask appropriate questions to ascertain:

- The type of information or records being requested
- The date span of the records being requested
- If appropriate, the method of sorting
- The desired physical media on which the records should be supplied
- The appropriate file format
- The appropriate supporting documentation necessary to make the requested records meaningful (data dictionaries, relationships within the data, field definitions, etc.)

Finally, the contact person should:

- Notify the requestor if any information or records s/he is requesting is exempt or otherwise not subject to the Public Records Act and therefore cannot be made available
- Inform the requestor of the estimated amount of time necessary to process and fulfill the request
- Make arrangements with the requestor for delivery of the requested records
- Arrange a method of communication between the requestor and the agency in the event that further information is needed by the agency to fulfill the request

EXEMPT RECORDS

The agency should be aware that some databases might be exempt from disclosure or contain information that should be redacted. For example, such information includes Social Security Numbers that are closed by Federal law. Section 149.43 of the Ohio Revised Code contains definitions of those records that are exempt from the Public Records Act. A public office has no obligation to make these records available to the

GUIDELINES FOR DATABASES AS PUBLIC RECORDS

OHIO ELECTRONIC RECORDS COMMITTEE

<http://www.ohiohistory.org/ohiojunction/erc/>

public and cannot be forced to make them available to the requestor. In the event that a database containing exempt information is requested, the agency should redact such exempt information while making the public information available to the requestor.

It is recommended that when databases are designed any exempt records should be contained in separate fields so that it can be easily identified and filtered out of a public records request.^{2 2}

REQUESTS FOR SPECIFIC RECORDS

Generally, most public records requests are made to obtain specific records from a database. In most instances, standard report queries can easily extract the records needed to satisfy the requestor. Where records already exist for purposes of the Public Records Act, your agency should have little problem in performing the necessary tasks to fulfill the request. However, in some cases programming may be necessary to create the records to fulfill the public records request. In these cases, the record does not already exist as a public record and the agency would have no obligation to fulfill the request. Nevertheless, the agency may choose to create a new record to fulfill the request.

REQUEST FOR ENTIRE DATABASES OR FOR PORTIONS OF DATABASES

Those who request an entire database or portions of a database should be notified of the potential problems that may occur if the information received is simply raw data without the necessary components of the entire database. Your agency should provide the data in the format in which it is maintained or any other format readily available to the agency without any additional programming. However, your agency should take reasonable measures to try to accommodate the requestor, including providing the requestor the raw data in the format desired if possible. When necessary, data should be exported in standardized computerized formats (for example, ASCII text delimited files or XML with appropriate schema).

In the event the records provided do not satisfy the requestor, or if it is impossible for your agency to provide the requestor the entire database, your agency should provide the requestor the reasons why the request cannot be fulfilled. These reasons may relate to legal limitations, logical database design, licensing limitations, security reasons, data links, proprietary software and hardware, etc.

In meeting requests for a portion of or an entire database, your agency should provide adequate information to assist the requestor in interpreting the information or records provided. Examples of the information you may provide includes, but is not limited to:

- The minimum software and hardware specifications your database requires
- Any unique and specialized proprietary software used in the system

GUIDELINES FOR DATABASES AS PUBLIC RECORDS

OHIO ELECTRONIC RECORDS COMMITTEE

<http://www.ohiohistory.org/ohiojunction/erc/>

- Data dictionaries
- Field definitions
- Relationships between tables

Whenever possible, records should be provided in standard and universally accepted formats. Your agency is not required to provide proprietary software or to provide software, hardware or any logical information that would compromise any system security or licensing regulations or agreements.

RECOMMENDATIONS

1. The Ohio Public Records Laws makes requests for databases a very real possibility. Keep this in mind when designing databases.
2. When designing a database, do not link exempt records with public records thus making it difficult to separate them when responding to a public records request.
3. Agencies, in order to facilitate public records requests are encouraged to consider logical database design, licensing limitations, security reasons, data links, proprietary software and hardware, etc. when designing databases.
4. The contact person should have easy access to descriptive information about the agency's databases, their contents and schemas in order to respond efficiently and accurately to a public records request.
5. Agencies with little or no experience with in-house database design are advised to seek out knowledgeable assistance when designing a database. This may be essential to achieving recommendations #2 and #3.

DEFINITIONS

- a. Column—A vertical list of fields from multiple records, a list from one field.
- b. Database—A database is a collection of data that is organized so that its contents can easily be accessed, managed, and updated. Databases contain aggregations of data records or files. The most prevalent type of database is the relational database, a tabular database in which data is defined so that it can be reorganized and accessed in a number of different ways. A distributed database is one that can be dispersed or replicated among different points in a network. An object-oriented programming database is one that is congruent with the data defined in object classes and subclasses.
- c. Data Set—see file
- d. Data definitions—Information regarding the layout, content, or use of a data field within a database. Examples of layout would include the type of file (text, excel version, HTML, ...), delimiters used (comma, tab, @, ...),

GUIDELINES FOR DATABASES AS PUBLIC RECORDS

OHIO ELECTRONIC RECORDS COMMITTEE

<http://www.ohiohistory.org/ohiojunction/erc/>

- field layout if fixed length (1-4 is field 1, 5-25 is field 2, ...), and other information the requestor would need to be able to load or use the data. Examples of content would include code tables when the output is not text (for field 15 1=Yes and 2=No, for field 12 1=Active, 2=Terminated, 3=Pending, ...), a list of the field names, definitions of the contents of the fields if not readily defined by the name, etc. Examples of the use of the output might include information on how to combine information when the output is in multiple files (use field 1 from file 1 = field 1 from file 2) so that the user can create their own query.
- e. Field—a defined area within a record. This definition includes a field name, a format (e.g. char, long, int.), and sometimes a length.
 - f. File—Two or more records of identical layout treated as a unit. The unit is larger than a record, but smaller than a data system, and is also known as a data set or file set.
 - g. Media—Physical storage media. A means of storing data. A piece of media allows data to be copied on to it, which can then be read back by a computer. Some types of media allow data to be recopied (destroying the original data in the process) while other types of media will only allow data to be copied to the media once. Common types of media are CD-ROM, magnetic tape, floppy disk, and paper.
 - h. Proprietary formats/software—Privately owned and controlled. In the computer industry, proprietary is the opposite of open. A proprietary design or technique is one that is owned by a company. It also implies that the company has not divulged specifications that would allow other companies to duplicate the product.
 - i. Query—A question, often required to be expressed in a formal way. In computers, what a user of a search engine or database enters is sometimes called a query. A database query can either be a select query or an action query. A select query is a data retrieval query. It specifies what fields/columns the user wants to retrieve as well as defines parameters/criteria that must be met for the data to be retrieved. Parameters can include date ranges, specific entries in a field/column, specific geographical regions, etc. The type and range of parameters will depend on the fields in the underlying data. Select queries can also include calculations on the data such as sum, minimum, maximum, etc. that act upon a specified field. An action query can ask for additional operations on the data, such as insertion, updating, or deletion.
 - j. Redact—Edit an image or document to render confidential information unreadable.
 - k. Relational Database—A relational database is a collection of data items organized as a set of formally-described tables from which data can be accessed or reassembled in many different ways without having to reorganize the database tables. The standard user and application program interface to a relational database is the structured query language (SQL).

GUIDELINES FOR DATABASES AS PUBLIC RECORDS

OHIO ELECTRONIC RECORDS COMMITTEE

<http://www.ohiohistory.org/ohiojunction/erc/>

SQL statements are used both for interactive queries for information from a relational database and for gathering data for reports. In addition to being relatively easy to create and access, a relational database has the important advantage of being easy to extend. After the original database creation, a new data category can be added without requiring that all existing applications be modified. A relational database is a set of tables containing data fitted into predefined categories. Each table (which is sometimes called a relation) contains one or more data categories in columns. Each row contains a unique instance of data for the categories defined by the columns. The definition of a relational database results in a table of metadata or formal descriptions of the tables, columns, domains, and constraints.

- l. Reports—Formatted output that takes its data from a query that was run against a database. Reports may include summary information and special formatting for the information displayed within the report.
- m. Row—In a relational database, a row consists of one set of attributes (or one tuple) corresponding to one instance of the entity that a table schema describes; a unique set of data for all the fields in a database
- n. Table—predefined format of rows and columns that define an entity.
- o. Tuple (database record)—A record is a collection of data items arranged for processing by a program. Multiple records are contained in a file or data set. The organization of data in the record is usually prescribed by the programming language that defines the record's organization and/or by the application that processes it. Typically, records can be of fixed-length or be of variable length with the length information contained within the record.

¹ The term "record(s)" as used in these guidelines refers to public records as defined in section 149.43 of the Ohio Revised Code. It is not used as a database or IT term.

² When creating databases with redacted layers, the file produced for distribution should be a combination of the original file and the redacted layer(s), so that they are one image with one layer. Once the new file is saved, it should not be possible to lift the redaction from the confidential fields.