ARCHAEOLOGICAL SITE DATABASE
DAVID DALY
SFSU ID: 913 694 774
GITHUB ID: DELTABDELT

Checkpoint #	Date Submitted
Checkpoint I	2024 Feb 19
Checkpoint I v2	
Checkpoint II	2024 Mar 3
Checkpoint II v2	
Checkpoint III	
Checkpoint III v2	
Checkpoint IV	
Checkpoint IV v2	
Checkpoint V	
Checkpoint V v2	

TABLE OF CONTENTS

Project Description

Archaeologists regularly attempt to balance two goals of the field: to learn more about humanity through material culture and physical traces left in the environment, and to protect and preserve cultural and historic resources, which includes forgoing excavation in the present to allow the possibility of more conscientious investigation with potentially more advanced technology in the future. It is critical, therefore, for archaeologists to record their observations of an archaeological site, whether as a result of field survey, geophysical investigation, shovel test pits, or excavation. Archaeologists also need the ability to deposit their findings in a central repository so that that information is (a) accessible to entities that need to use it (such as other archaeologists or governmental agencies who are voting on a development project) and also (b) protected from people who should not have it (such as bottle collectors or potential looters).

Information about archaeological sites in the state of California is housed in the California Historic Resources Information Service (CHRIS), which is overseen by the state Office of Historic Preservation (OHP). This data is then made available to qualified individuals via nine Information Centers, located at universities around the state (CSUs Bakersfield, Chico, Fullerton, Sacramento, and Stanislaus; Sonoma State University; the Santa Barbara Museum of Natural History; and UC Riverside).

I have extensive background in archaeology. I have a master's degree from San Francisco State in anthropology and archaeology. I also have years of experience in archaeology, collecting data in the field, preparing and submitting site record forms, and requesting and analyzing site records from Information Centers.

Although the structure of the CHRIS database system is not publicized, it is clear that the database records must contain fields that are part of the standardized site record forms, created and maintained by the Office of Historic Preservation (https://ohp.parks.ca.gov/?page_id=28351). These forms are typically referred to as "DPR forms" because they are published by the California Department of Parks and Recreation. The principal unit of the database is the site, denoted by a site identifier of the form P-XX-YYYYYY, where the two-digit number XX corresponds to the alphabetical order of the county in which the site is recorded. The second, six-digit number YYYYYY corresponds to the ordinal of the site's recording in that county.

Each site will have at least one corresponding site record, which is the most basic means of recording the site. Within the primary site record will be any other identifiers of the site (such as a site name or descriptor). The primary record also has a choice for whether the location of the site is "Not for Publication" or "Unrestricted": the latter category usually applies to well-known

structures such as the Golden Gate Bridge, the Stanford University Quad, or a city's historic district, while the former applies to the majority of archaeological deposits — from obsidian stone tool work areas to gold rush era trash pits to human burials, most archaeological site locations are not publicly available. The primary record also requires the location of the site (both identifying the county and the corresponding U.S. Geological Survey quadrangle map), as well as a text description of the resource. The resources present are listed by choosing corresponding codes from a list provided by OHP. The identity of the property owner is recorded, as is the name and affiliation of the recorder, the date recorded, the type of survey (for example, pedestrian survey, or excavation), and any reports associated with the current recordation.

Other DPR forms provide specialized fields that may or may not apply to any given site (such as rock art or historic districts). Each site is associated with a digital location, in the form of an ArcGIS shape file.

Functional Database Requirements

1. Site (Strong)
   1.1. A site shall be associated with multiple site records
   1.2. A site shall have a site identifier
   1.3. A site shall have several counties
   1.4. A site may consist of other sites (in a district)
   1.5.
2. Site record (Strong)
   2.1. A site record shall be associated with multiple sites
   2.2. A site record shall have multiple site identifiers
   2.3. A site record shall have a privacy setting
   2.4. A site record shall be associated with multiple counties
   2.5. A site record shall have a USGS quad
   2.6. A site record shall have a site description
   2.7. A site record shall have many resource codes
   2.8. A site record shall have many resource types
   2.9. A site record shall have an age and source
   2.10. A site record shall have an owner
   2.11. A site record shall have a recorder
   2.12. A site record shall have several addresses
   2.13. A site record shall have multiple survey types
   2.14. A site record shall have many report citations
   2.15. A site record shall have many attachments
   2.16. A site record shall have many maps
   2.17. A site record shall have a shape file
   2.18. A site record shall have one or more photos
3. Site identifier
   3.1. A site identifier shall pertain to a site
   3.2. A site identifier shall have at least one site record
   3.3. A site identifier shall be of the form “P-##-######”
4. County (Strong)
   4.1. A county shall have many sites
   4.2. A county shall have many site records
   4.3. A county shall be associated with one information center

5. USGS quad (Strong)
   5.1. A USGS quad may have many site records
   5.2. A USGS quad shall be the name of a valid USGS quad map
6. Site description
   6.1. A site description shall pertain to one site record
7. Resource code (Strong)
   7.1. A resource code shall pertain to many site records
8. Resource type (Strong)
   8.1. A resource type shall pertain to many site records
9. Owner (Strong)
   9.1. An owner may pertain to many site records
   9.2. An owner may have several addresses
10. Recorder (Strong)
   10.1. A recorder may have multiple affiliations
   10.2. A recorder may have multiple addresses
11. Affiliation (Strong)
   11.1. An affiliation may pertain to many recorders
   11.2. An affiliation may pertain to many users
12. Address (Strong)
   12.1. An address may pertain to many owners
   12.2. An address may pertain to many recorders
   12.3. An address shall pertain to many information centers
13. Survey type (Strong)
   13.1. A survey type may pertain to several site records
14. Report citation (Strong)
   14.1. A report citation may pertain to multiple site records
   14.2. A report citation may be “None”
15. Attachments (Strong)
   15.1. An attachment shall pertain to one or many site records
   15.2. An attachment may be “None”
16. Map (Strong)
   16.1. A map shall pertain to several site records
   16.2. A map may be a sketch map or a location map
17. Location map (Weak)
   17.1. A location map is a type of map

18. Sketch map (Weak)
    18.1.A sketch map is a type of map
19. Shape file
    19.1. A shape file shall pertain to one site record
20. Photo
    20.1. A photo shall pertain to one site record

21. User (Strong)
    21.1. A user shall have an account
    21.2. A user shall have many affiliations
    21.3.A user may work for at most one information center at a time
22. Account
    22.1. An account shall pertain to only one user
    22.2. An account shall have a password
    22.3. An account may engage in at most one session at a time
23. Session
    23.1. A session shall belong to only one account
24. Information center
    24.1. An information center shall serve many counties
    24.2. An information center shall employ many users
    24.3. An information center shall have an address

Non-functional Database Requirements

1. Performance
   1.1. The database system shall support concurrent queries.
   1.2. Because updates and queries can be scheduled so as not to overlap, support for concurrent updates is not a priority.
   1.3. The database system will need to deliver ArcGIS shape files and images reliably.
   1.4. The database shall be optimized for efficient queries.
   1.5. The system should be configured for reliability.

2. Security
   2.1. The database system shall allow access to authorized users.
   2.2. The database system shall not allow access to anyone who is not an authorized user.
   2.3. The database system shall support only encrypted passwords.
   2.4. The database system shall accept only data consistent with the corresponding domain.
   2.5. The database shall not support more than one session for a given user.

3. Scalability
   3.1. Because the number of authorized users, either at an Information Center or in a cultural resource management firm, is small, the system should begin small.
   3.2. Performance checks should be run to ensure throughput is sufficient.
   3.3. Horizontal scaling is possible if staffing increases.
   3.4. The likelihood of substantial horizontal scaling is miniscule.
   3.5. DPR forms were last updated in 1993. This stability means that functional scalability is not a high priority.

4. Capability
   4.1. The system shall offer acceptable responses when check boxes appear on DPR forms.
   4.2. The system shall check the format of the primary identifier field.
   4.3. The system shall validate formatting of date fields.
   4.4. The system shall inform the user of allowable file formats.
   4.5.The system shall inform the user of appropriate data formatting.

5. Environmental
   5.1. The system will be used in office settings.
   5.2. The system will be used on desktop computers.
   5.3. The system should not have a mobile interface.
   5.4. The system does not require touch-screen capability.
   5.5.The system should provide legible, useful data formatting.

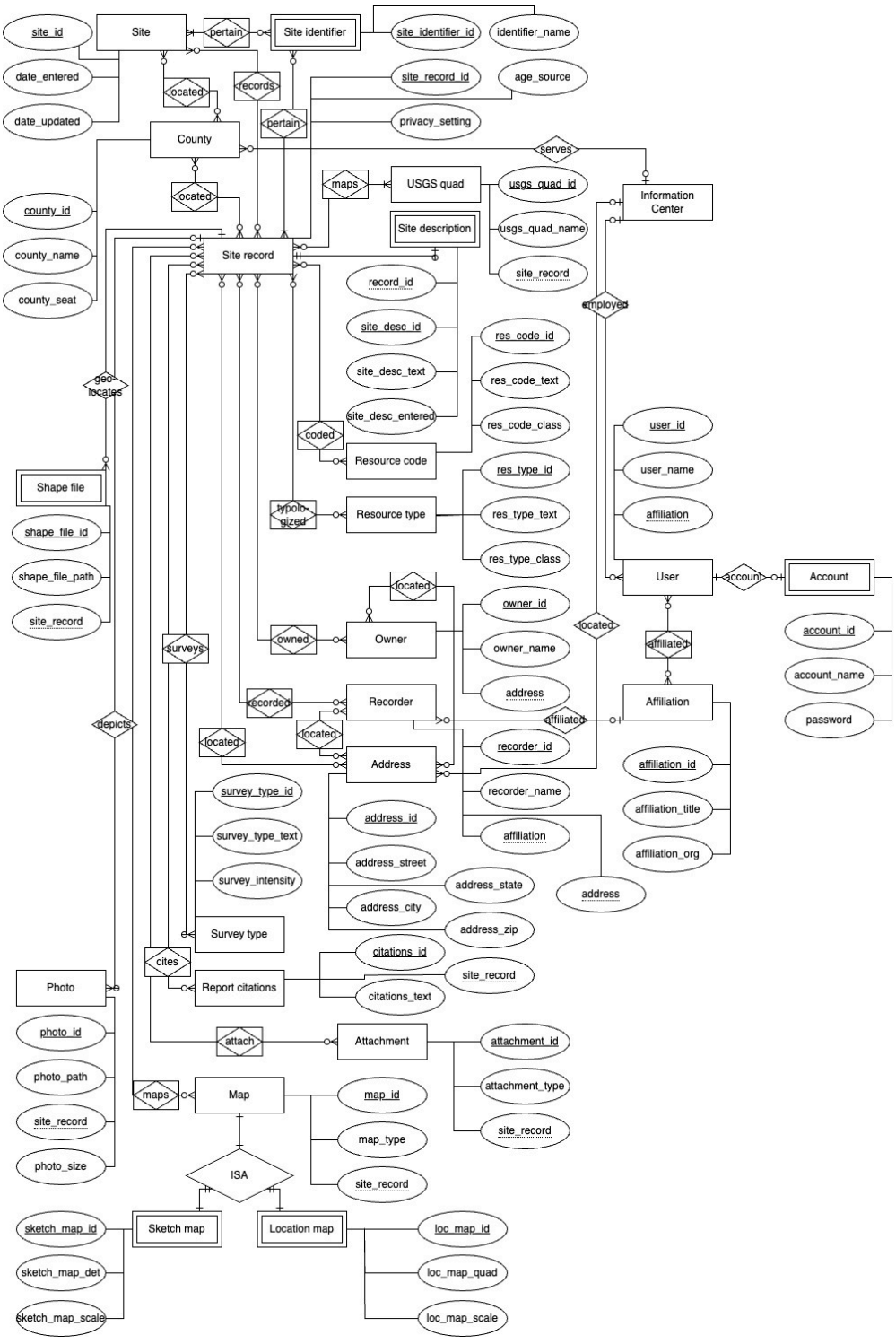
6. Coding Standards
   6.1. The system should be compatible with major operating systems.
   6.2. If the system cannot be compatible with both Windows and Mac, it should be compatible with Windows.
   6.3. The system should support the current version of mySQL.
   6.4. The system should not be needlessly complicated, so that it can be compatible with more versions of mySQL.
   6.5. The system should be accompanied with documentation for usability.

7. Media Storage
   7.1. The system shall store ArcGIS shape files in an appropriate format.
   7.2. The system shall store images as JPG or TIF files.
   7.3. The system shall store digital versions of site records as PDF files.
   7.4. The system shall store digital versions of reports as PDF files.

7.5. The system shall not store files such as EXE files.

8. Privacy
   8.1. The database system shall not ask for sensitive personal information of users.
   8.2. The database shall maintain names and addresses of property owners.
   8.3. The database shall maintain names, addresses, and affiliations of site recorders.
   8.4.The database shall maintain the precise location of archaeologically sensitive areas, including sites that contain human remains.
   8.5. The database shall not release aforementioned data, except to an approved user.

Entity Relationship Diagram (ERD)



Entity Description

List of Entities and Attributes

1. Site (Strong)
   1.1. site_id: key, numeric
   1.2. date_entered: multivalue, timestamp
   1.3. date_updated: multivalue, timestamp
2. Site record (Strong)
   2.1. site_record_id: key, numeric
   2.2. privacy_setting: numeric (“Not for Publication” or “Unrestricted”)
   2.3. age_source: numeric (“Historic”, “Prehistoric”, or “Both”)
3. Site identifier (Weak)
   3.1. site_identifier_id: key, numeric
   3.2. identifier_name: alphanumeric
   3.3. site: key, numeric
   3.4. site_record: key, numeric
4. County (Strong)
   4.1. county_id: key, numeric
   4.2. county_name: alphanumeric
   4.3. county_seat: alphanumeric
5. USGS quad (Strong)
   5.1. usgs_quad_id: key, numeric
   5.2. usgs_quad_name: alphanumeric
   5.3. site_record: key, numeric
6. Site description (Weak)
   6.1. site_desc_id: key, numeric
   6.2. site_desc_text: alphanumeric
   6.3. site_desc_entered: multivalue, timestamp
7. Resource code (Strong)
   7.1. res_code_id: key, numeric
   7.2. res_code_text: alphanumeric
   7.3. res_code_class: alphanumeric
8. Resource type (Strong)
   8.1. res_type_id: key, numeric

8.2. res_type_name: alphanumeric
8.3. res_type_desc: alphanumeric
9. Owner (Strong)
9.1. owner_id: key, numeric
9.2. owner_name: alphanumeric
9.3. address: key, numeric
10.Recorder (Strong)
10.1. recorder_id: key, numeric
10.2. recorder_name: alphanumeric
10.3. affiliation: key, numeric
10.4. address: key, numeric
11.Affiliation (Strong)
11.1. affiliation_id: key, numeric
11.2. affiliation_org: alphanumeric
11.3. affiliation_title: alphanumeric
12.Address (Strong)
12.1. address_id: key, numeric
12.2. address_street: alphanumeric
12.3. address_city: alphanumeric
12.4. address_state: alphanumeric
12.5. address_zip: alphanumeric
13.Survey type (Strong)
13.1.survey_type_id: key, numeric
13.2. survey_type_text: alphanumeric
13.3. survey_intensity: numeric
14.Report citations (Strong)
14.1. citations_id: key, numeric
14.2. citations_text: alphanumeric
14.3. site_record: key, numeric
15.Attachment (Strong)
15.1. attachment_id: key, numeric
15.2. attachment_type: alphanumeric
15.3. site_record: key, numeric
16. Map (Strong)
16.1. map_id: key, numeric
16.2.map_type: numeric

16.3. site_record: key, numeric
17.Location map (Weak)
17.1. loc_map_id: key, numeric
17.2. loc_map_quad: alphanumeric
17.3. loc_map_scale: numeric
18.Sketch map (Weak)
18.1. sketch_map_id: key, numeric
18.2. sketch_map_quad: alphanumeric
18.3. sketch_map_scale: numeric
19.Photo (Strong)
19.1. photo_id: key, numeric
19.2. photo_path: alphanumeric
19.3. site_record: key, numeric
19.4. photo_size: numeric
20.Shape file (Strong)
20.1. shape_file_id: key, numeric
20.2. shape_file_path: alphanumeric
20.3. site_record: key, numeric
21.User (Strong)
21.1. user_id: key, numeric
21.2. user_name: alphanumeric
21.3. affiliation: key, numeric
22.Account (Weak)
22.1. account_id: key, numeric
22.2. account_name: alphanumeric
22.3. password: alphanumeric
23. Session (Weak)
23.1. session_id: key, numeric
23.2. account_id: numeric
24. Information center (Strong)
24.1. info_center_id: key, numeric
24.2. county_id: numeric
24.3. user_id: numeric
24.4. address_id: numeric

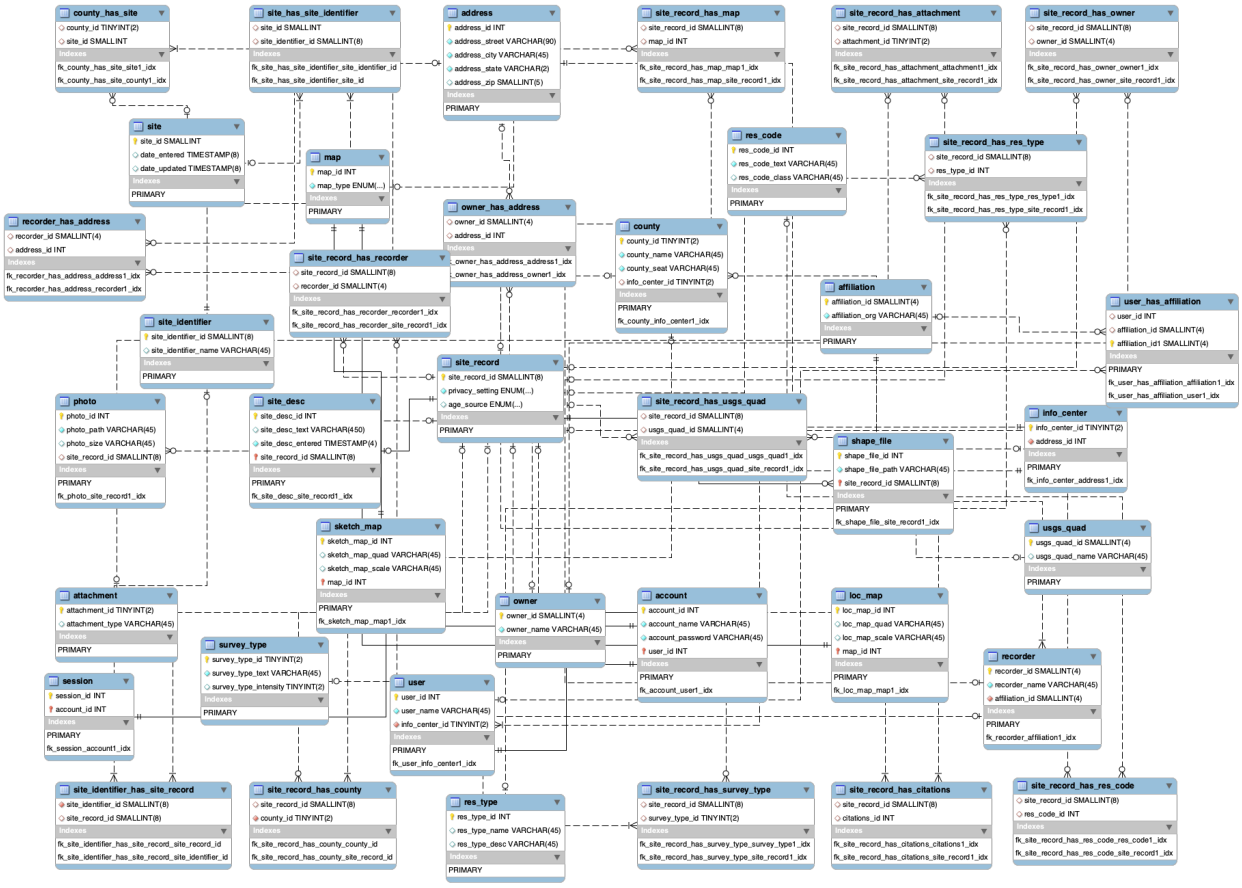
Associative entities

1. site_identifier_site
    1.1. site_identifier_site_id: key, numeric
    1.2. site_id: key, numeric
    1.3. site_identifier_id: key, numeric
2. site_county
    2.1. site_county_id: key, numeric
    2.2. site_id: key, numeric
    2.3. county_id: key, numeric
3. site_record_site
    3.1. site_record_site_id: key, numeric
    3.2. site_id: key, numeric
    3.3. site_record_id: key, numeric
4. site_identifier_site_record
    4.1. site_identifier_site_record_id: key, numeric
    4.2. site_id: key, numeric
    4.3. site_record_id: key, numeric
5. site_record_county
    5.1. site_identifier_county_id: key, numeric
    5.2. county_id: key, numeric
    5.3. site_record_id: key, numeric
6. site_record_usgs_quad
    6.1. site_record_usgs_quad_id: key, numeric
    6.2. usgs_quad_id: key, numeric
    6.3. site_record_id: key, numeric
7. site_record_res_code
    7.1. site_record_res_code_id: key, numeric
    7.2. res_code_id: key, numeric
    7.3. site_record_id: key, numeric
8. site_record_res_type
    8.1. site_record_res_type_id: key, numeric
    8.2. res_type_id: key, numeric
    8.3. site_record_id: key, numeric
9. site_record_owner
    9.1. site_record_owner_id: key, numeric
    9.2. owner_id: key, numeric

9.3. site_record_id: key, numeric
10. site_record_recorder
    10.1. site_record_recorder_id: key, numeric
    10.2. recorder_id: key, numeric
    10.3. site_record_id: key, numeric
11. site_record_address
    11.1. site_record_address_id: key, numeric
    11.2. address_id: key, numeric
    11.3. site_record_id: key, numeric
12. site_record_survey_type
    12.1. site_record_survey_type_id: key, numeric
    12.2. survey_type_id: key, numeric
    12.3. site_record_id: key, numeric
13. site_record_citations
    13.1. site_record_citations_id: key, numeric
    13.2. citations_id: key, numeric
    13.3. site_record_id: key, numeric
14. site_record_map
    14.1. site_record_map_id: key, numeric
    14.2. map_id: key, numeric
    14.3. site_record_id: key, numeric
15. site_record_attachment
    15.1. site_record_attachment_id: key, numeric
    15.2. attachment_id: key, numeric
    15.3. site_record_id: key, numeric
16. user_affiliation
    16.1. user_affiliation_id: key, numeric
    16.2. user_id: key, numeric
    16.3. affiliation_id: key, numeric
17. owner_address
    17.1. owner_address_id: key, numeric
    17.2. owner_id: key, numeric
    17.3. address_id: key, numeric
18. recorder_address
    18.1. recorder_address_id: key, numeric
    18.2. recorder: key, numeric

18.3. address_id: key, numeric

Entity Establishment Relationship Diagram (EER)



Constraints Description

Table	FK	ON DELETE	ON UPDATE	Comment
site	site_id entifier	SET NULL	CASCADE	A site may have many site identifiers and should remain if a site identifier is deleted
site_id entifier	site	SET NULL	CASCADE	A site identifier is also tied to site records and should remain if a site is deleted
site_ record	site_id entifier	SET NULL	CASCADE	A site record may have many site identifiers and should remain if a site identifier is deleted
site_id entifier	site_ record	SET NULL	CASCADE	A site identifier may be tied to other site records and should remain if a site record is deleted
site	county	SET NULL	CASCADE	A site should remain for reassignment if a county is deleted
county	site	SET NULL	CASCADE	A county may be tied to other sites and should remain if a site is deleted
site_ record	county	SET NULL	CASCADE	A site record should remain for reassignment if a county is deleted
county	site_ record	SET NULL	CASCADE	A county may be tied to other sites record and should remain if a site record is deleted
site_ record	usgs_ quad	SET NULL	CASCADE	A site record should remain for reassignment if a USGS quad is deleted
usgs_ quad	site_ record	SET NULL	CASCADE	A USGS quad may be tied to other sites record and should remain if a site record is deleted
site_ desc	site_ record	CASCADE	CASCADE	The site description is tied to the site record; if deleted or modified, changes should cascade

site_ record	res_ code	SET NULL	CASCADE	A site record should remain for reassignment if a resource code is deleted
res_ code	site_ record	SET NULL	CASCADE	A resource code may be tied to other site records and should remain if a site record is deleted
site_ record	res_ type	SET NULL	CASCADE	A site record should remain for reassignment if a resource type is deleted
res_ type	site_ record	SET NULL	CASCADE	A resource type may be tied to other site records and should remain if a site record is deleted
site_ record	owner	SET NULL	CASCADE	A site record should remain for reassignment if an owner is deleted
owner	site_ record	SET NULL	CASCADE	An owner may be tied to other site records and should remain if a site record is deleted
owner	addres s	SET NULL	CASCADE	An address should remain for reassignment if an owner is deleted
addres s	owner	SET NULL	CASCADE	An owner may be tied to other addresses and should remain if an address is deleted
site_ record	recorde r	SET NULL	CASCADE	A site record should remain for reassignment if a recorder is deleted
recorde r	site_ record	SET NULL	CASCADE	A recorder may be tied to other site records and should remain if a site record is deleted
recorde r	addres s	SET NULL	CASCADE	An address should remain for reassignment if a recorder is deleted
addres s	recorde r	SET NULL	CASCADE	A recorder may be tied to other addresses and should remain if an address is deleted
recorde r	affiliatio n	SET NULL	CASCADE	An affiliation may be tied to other recorders and should remain if a recorder is deleted
user	affiliatio n	SET NULL	CASCADE	An affiliation should remain for reassignment if a user is deleted

affiliation	user	SET NULL	CASCADE	A user may be tied to other affiliation and should remain if an affiliation is deleted
account	user	CASCADE	CASCADE	If a user is deleted, the user's account should also be deleted
user	info_center	SET NULL	CASCADE	Deleting an info center should free the former employees for reassignment, rather than deleting
info_center	address	NO ACTION	CASCADE	Deleting an info center should not affect the address, which may be associated with other fields
info_center	county	NO ACTION	CASCADE	Deleting an info center should not affect the county, which may be associated with other fields
site_record	survey_type	SET NULL	CASCADE	A site record should remain for reassignment if a survey type is deleted
survey_type	site_record	SET NULL	CASCADE	A survey type may be tied to other site records and should remain if a site record is deleted
site_record	Citations	SET NULL	CASCADE	A site record should remain for reassignment if a citation is deleted
Citations	site_record	SET NULL	CASCADE	A citation may be tied to other site records and should remain if a site record is deleted
site_record	attachment	SET NULL	CASCADE	A site record should remain for reassignment if an attachment is deleted
attachment	site_record	SET NULL	CASCADE	An attachment may be tied to other site records and should remain if a site record is deleted
site_record	map	SET NULL	CASCADE	A site record should remain for reassignment if a map is deleted
map	site_record	SET NULL	CASCADE	A map may be tied to other site records and should remain if a site record is deleted

photo	site_ record	CASCADE	CASCADE	If a site record is deleted, the site record's photo(s) should also be deleted, because of the one-to-many relationship
shape_ file	site_ record	CASCADE	CASCADE	If a site record is deleted, the site record's shape file should also be deleted, because of the one-to-many relationship
session	accoun t	CASCADE	CASCADE	If an account is deleted, any associated session should also be deleted, because of the one-to-one relationship
loc_ map	map	CASCADE	CASCADE	If a location map is deleted, the map should also be deleted, because of the one-to-one relationship
sketch_ map	map	CASCADE	CASCADE	If a sketch map is deleted, the map should also be deleted, because of the one-to-one relationship