

New York State Physician Profile Database Manual

March 2009

PROPRIETARY

NEW YORK STATE PHYSICIAN PROFILE DATABASES AND LINKAGES MANUAL

This manual provides the technical details of the design of the data model and its attributes (specific data elements) and tables (groups of attributes) as well as the relationships between these attributes and tables. Additionally, this manual provides detailed source and usage information for data attributes. Specifically, external attributes are identified with their primary and secondary sources and if physicians are allowed to change the database information.

This manual is organized as follows:

Section 1: Data Model

Section 2: Data Dictionary

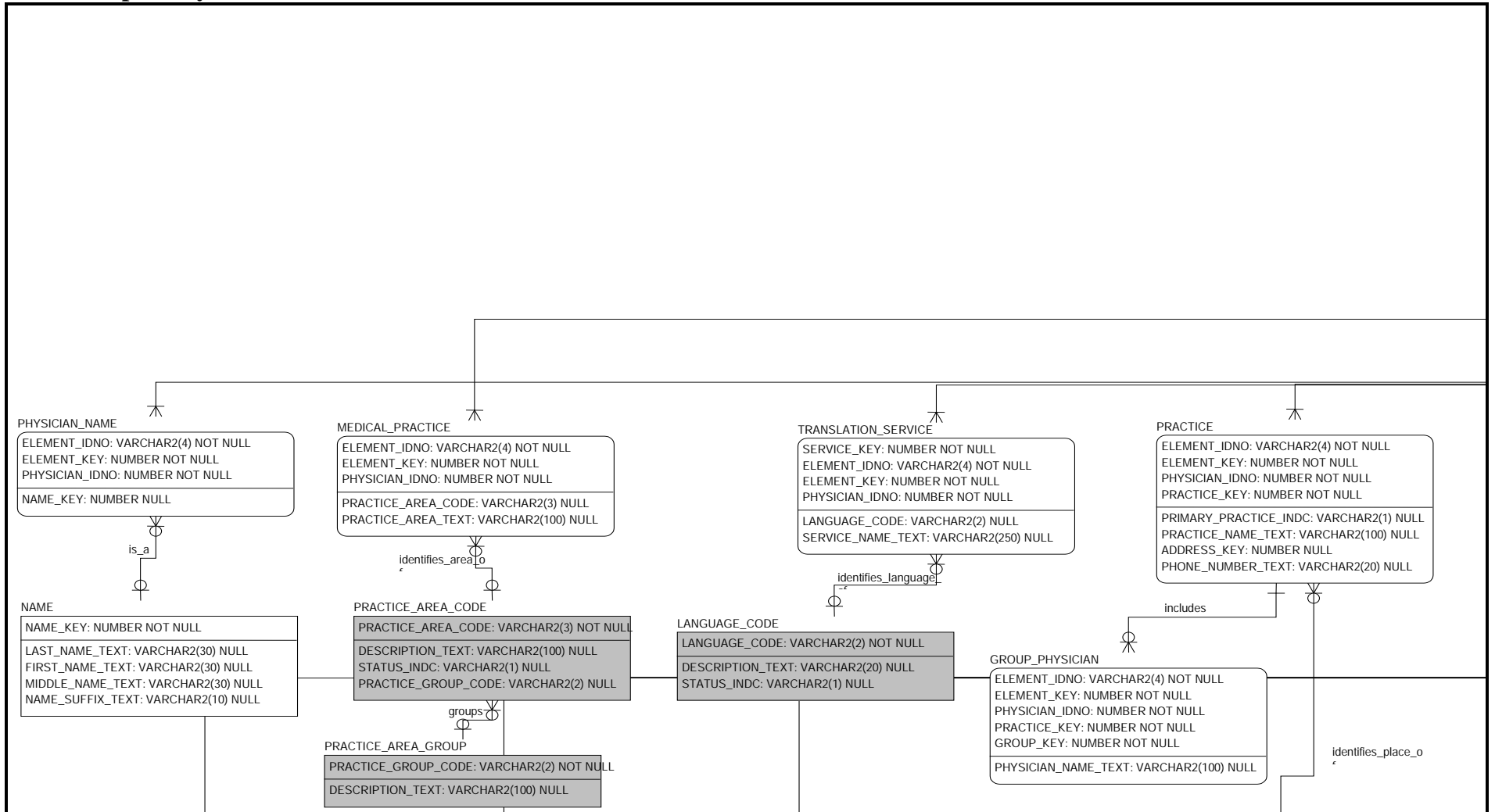
Section 3: External Files and Import Procedures

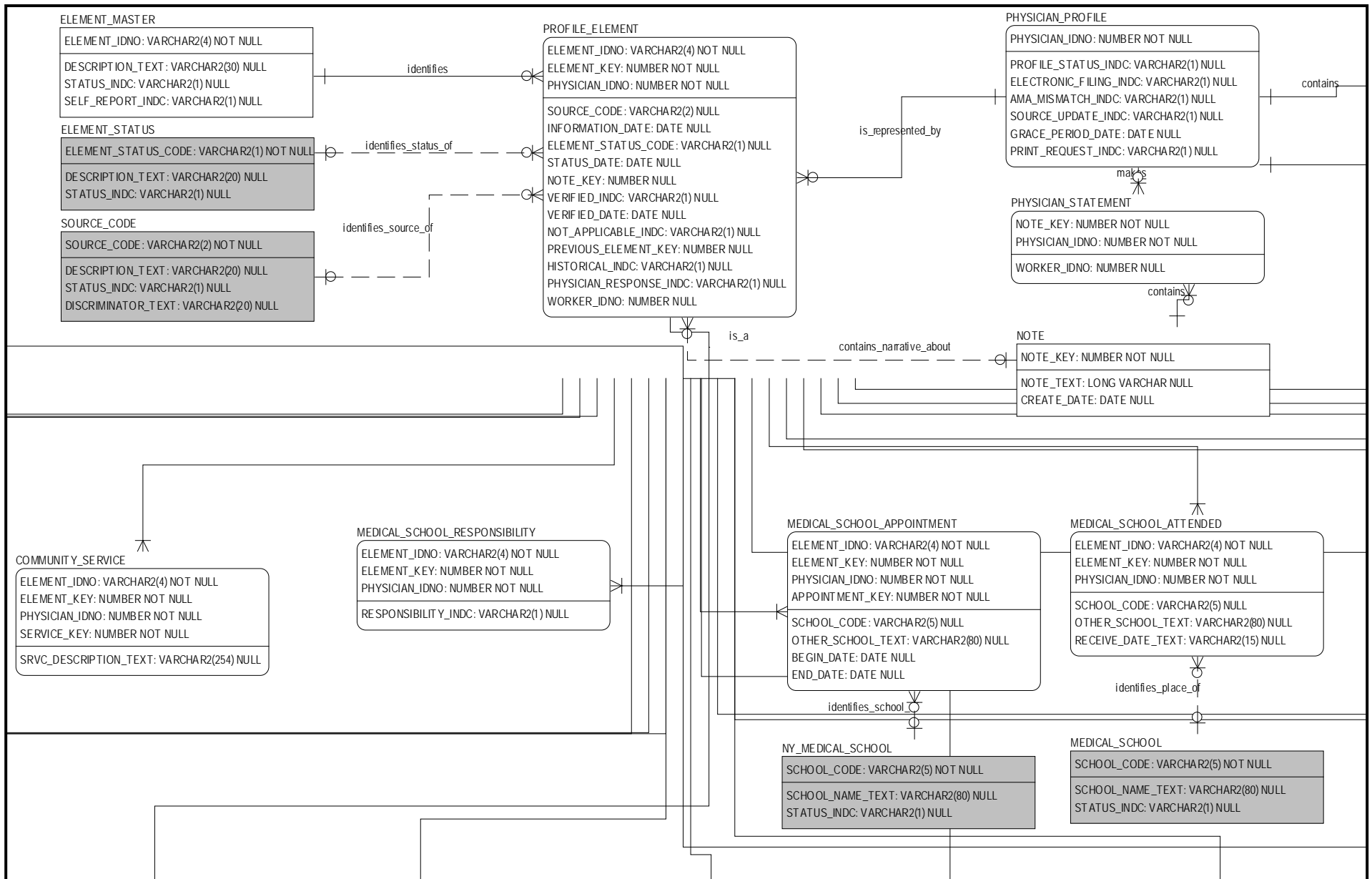
1. Data Model

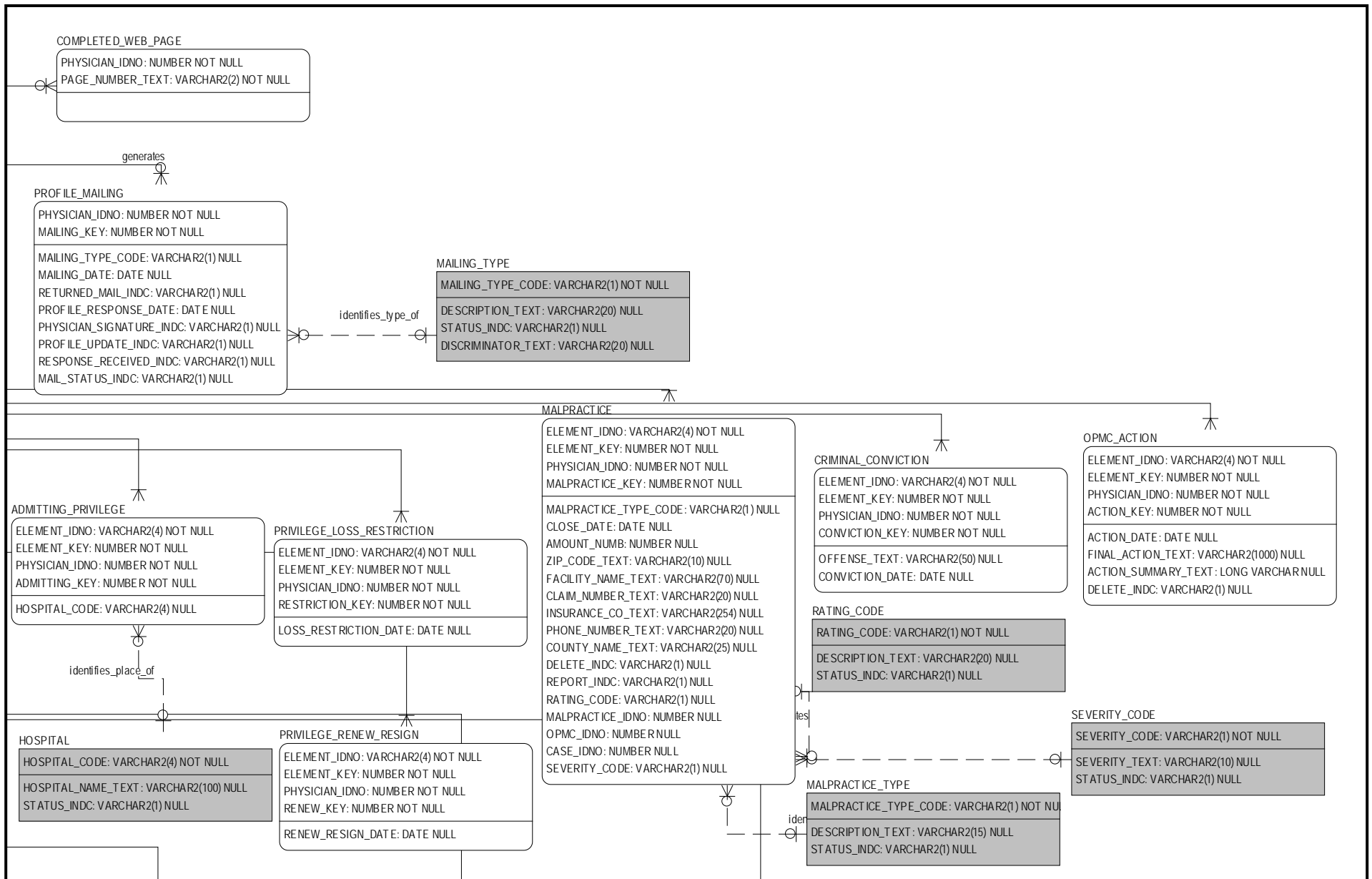
The New York State Physician Profile Data Model provides a graphical representation of the relationships between the various tables and attributes used in the New York State Physician Profile data system. The data model has two separate and distinct components, the Central Repository and The Public Web. Each of these is graphically displayed on the pages that follows:

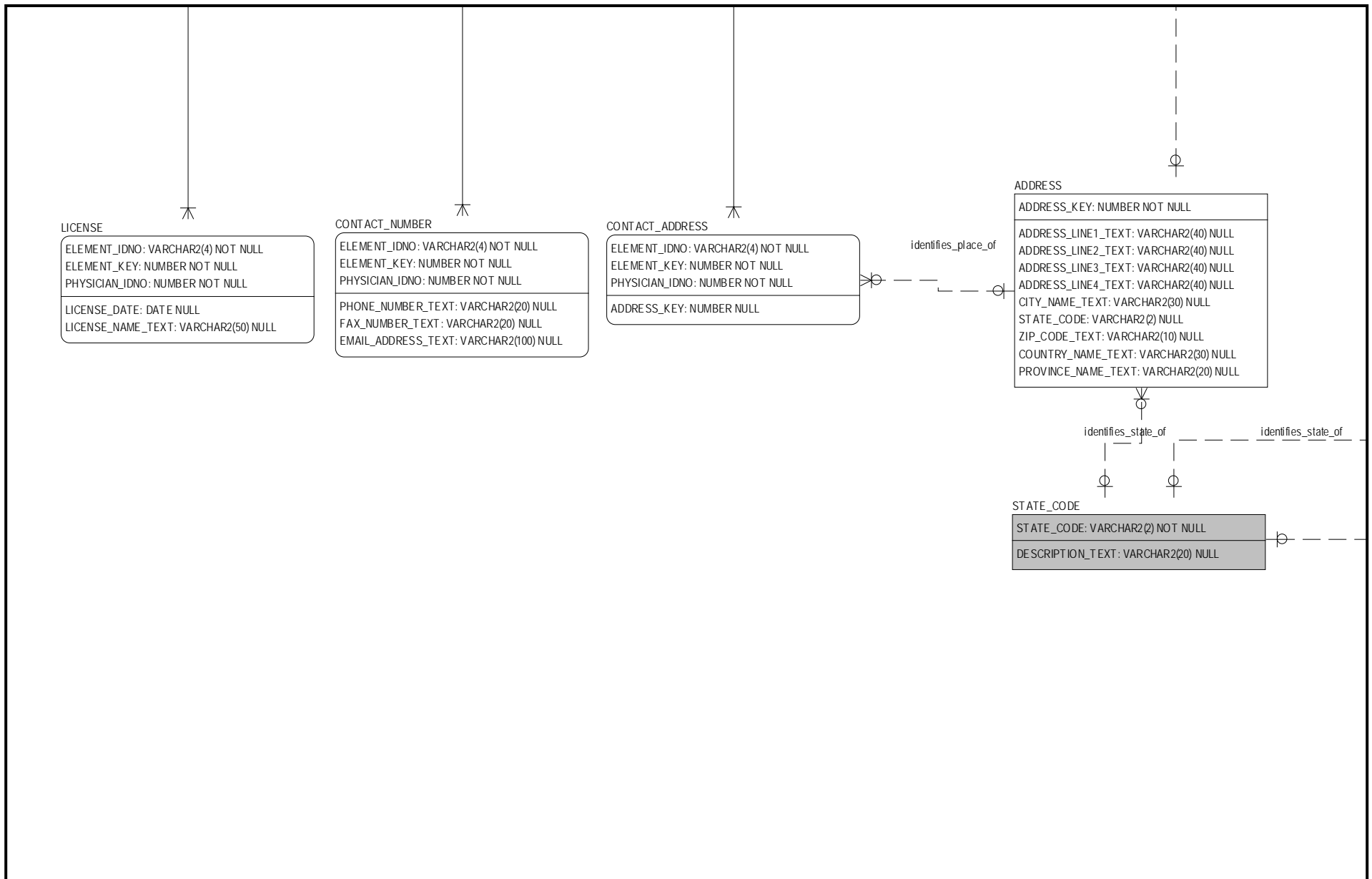
NY PHYSICIAN PROFILE DATA MODEL

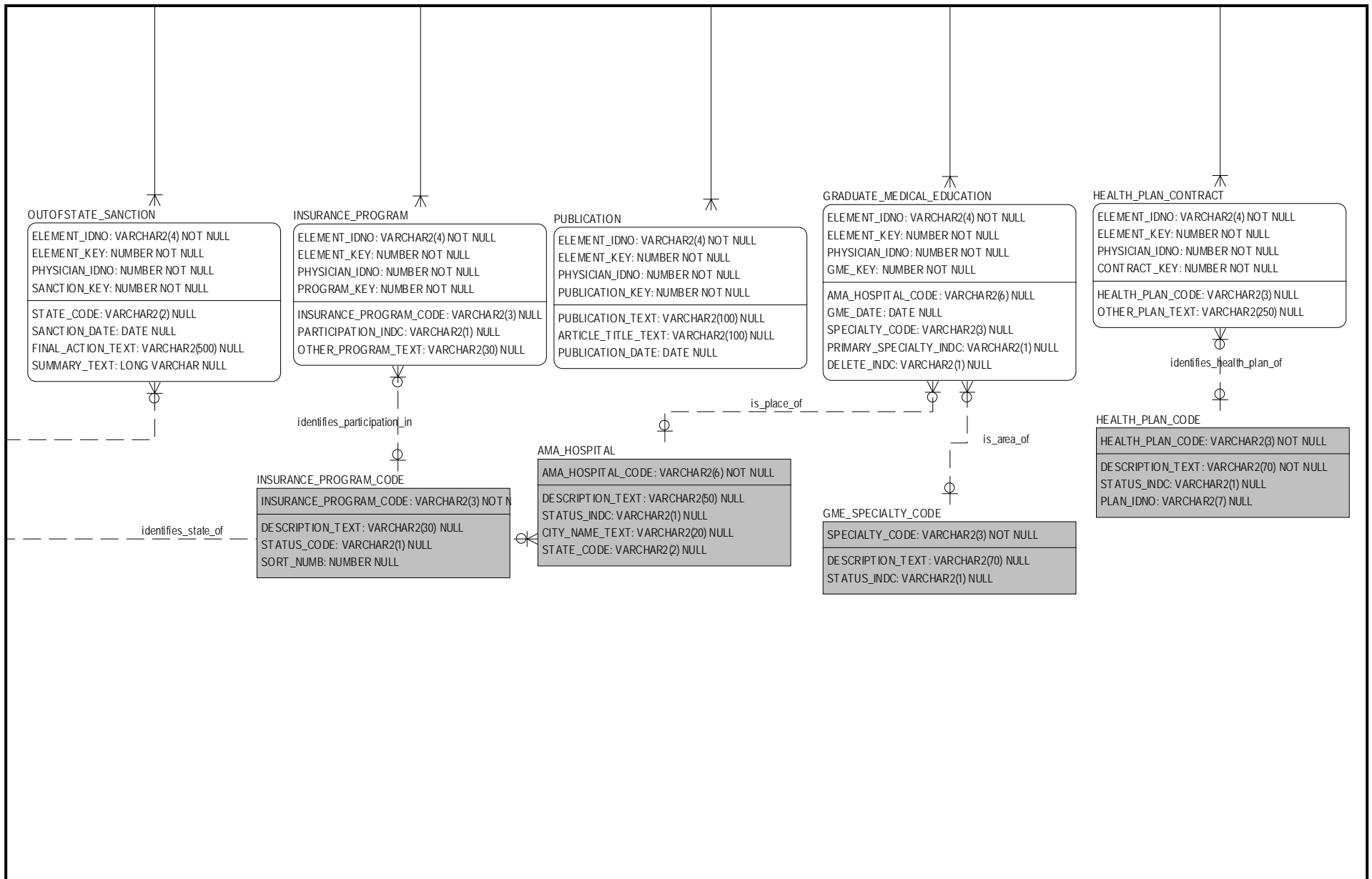
Central Repository Elements

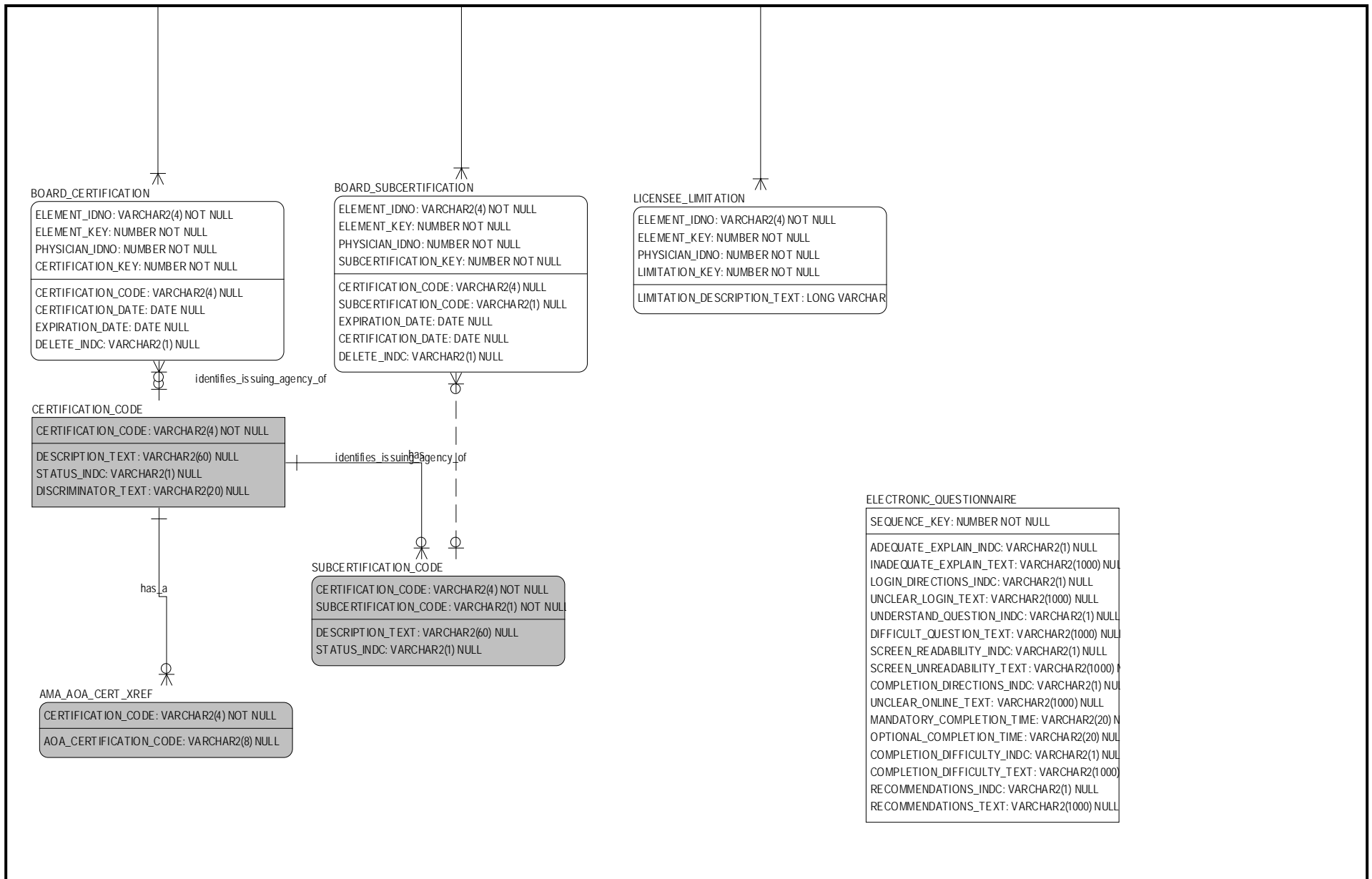




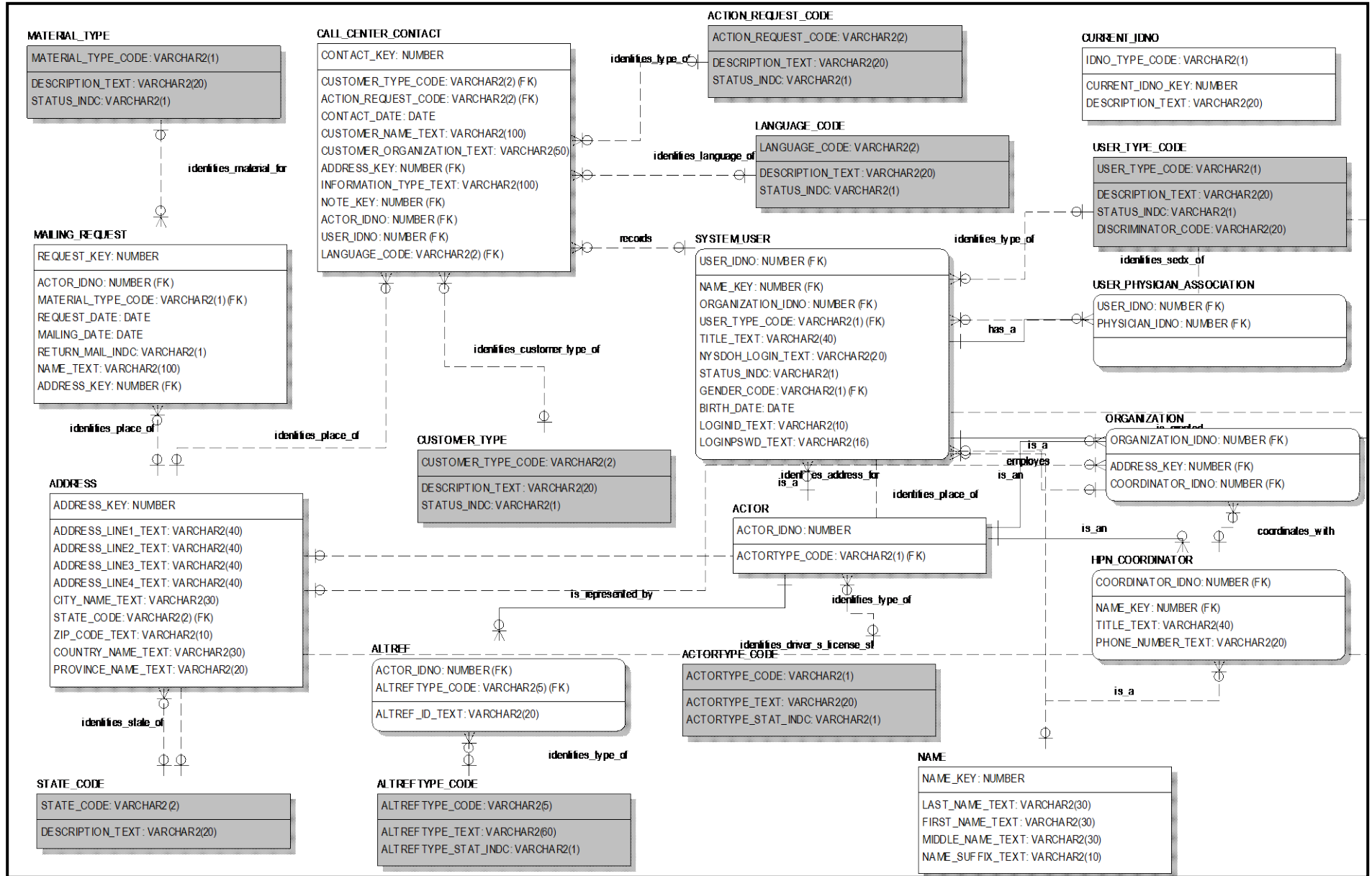








CALL CENTER/SECURITY



MISCELLANEOUS WEB PUBLICATION

LICENSEE_LIMITATION_STAGING

ELEMENT_IDNO: VARCHAR2(4)
ELEMENT_KEY: NUMBER
PHYSICIAN_IDNO: NUMBER
LIMITATION_KEY: NUMBER

LIMITATION_DESCRIPTION_TEXT: CLOB
INFORMATION_DATE: DATE

PHYSICIAN_STATEMENT_STAGING

NOTE_KEY: NUMBER
PHYSICIAN_IDNO: NUMBER

NOTE_TEXT: CLOB
CREATE_DATE: DATE

PUBLIC_WEB_FILTER

PHYSICIAN_IDNO: NUMBER

FILTER_TYPE_CODE: VARCHAR2(1)

MEDSCHOOL_RESP_STAGING

ELEMENT_IDNO: VARCHAR2(4)
ELEMENT_KEY: NUMBER
PHYSICIAN_IDNO: NUMBER

RESPONSIBILITY_INDC: VARCHAR2(1)
INFORMATION_DATE: DATE

OUTOFSTATE_SANCTION_STAGING

ELEMENT_IDNO: VARCHAR2(4)
ELEMENT_KEY: NUMBER
PHYSICIAN_IDNO: NUMBER
SANCTION_KEY: NUMBER

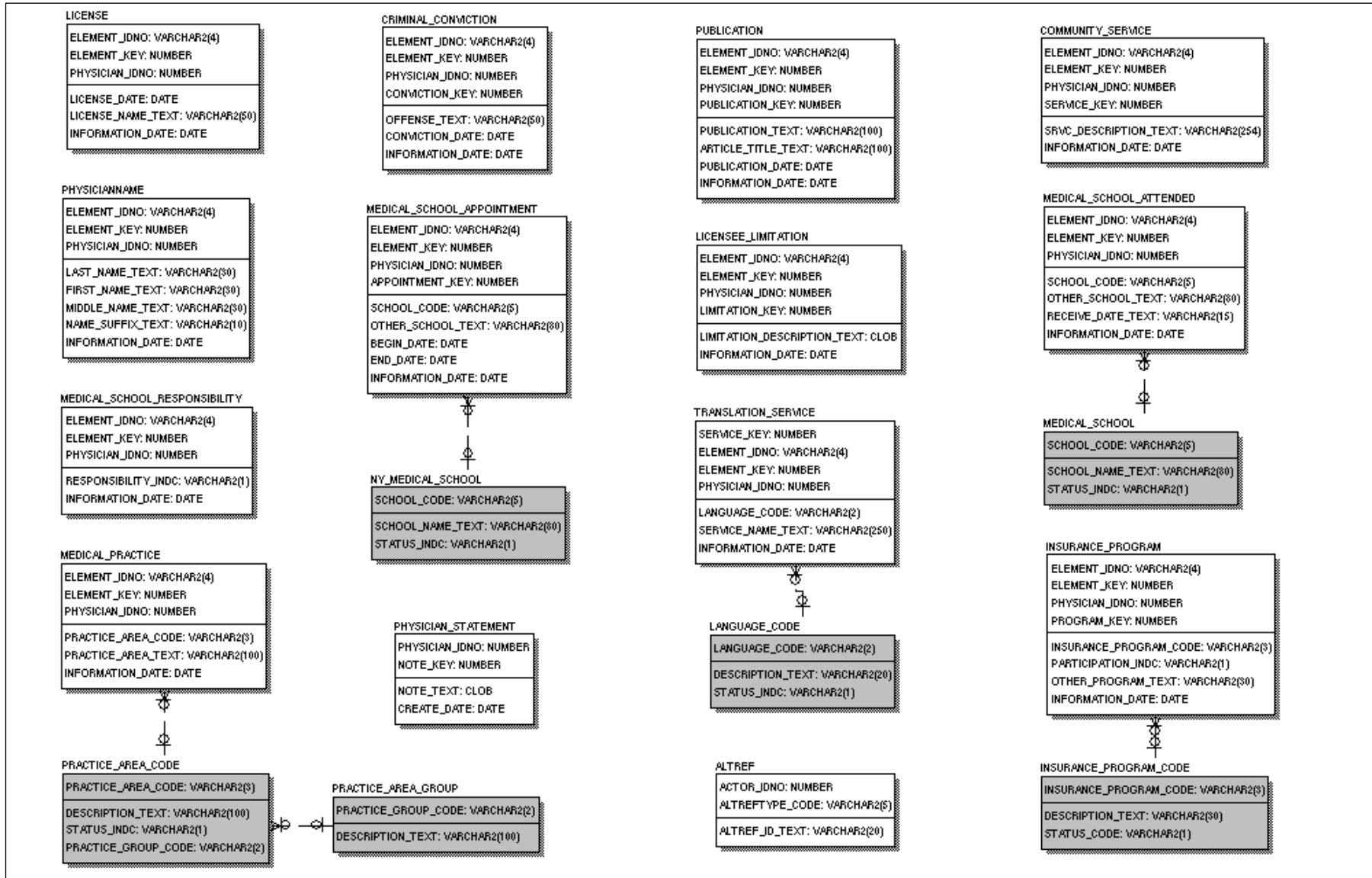
STATE_CODE: VARCHAR2(2)
SANCTION_DATE: DATE
FINAL_ACTION_TEXT: VARCHAR2(500)
SUMMARY_TEXT: CLOB
INFORMATION_DATE: DATE

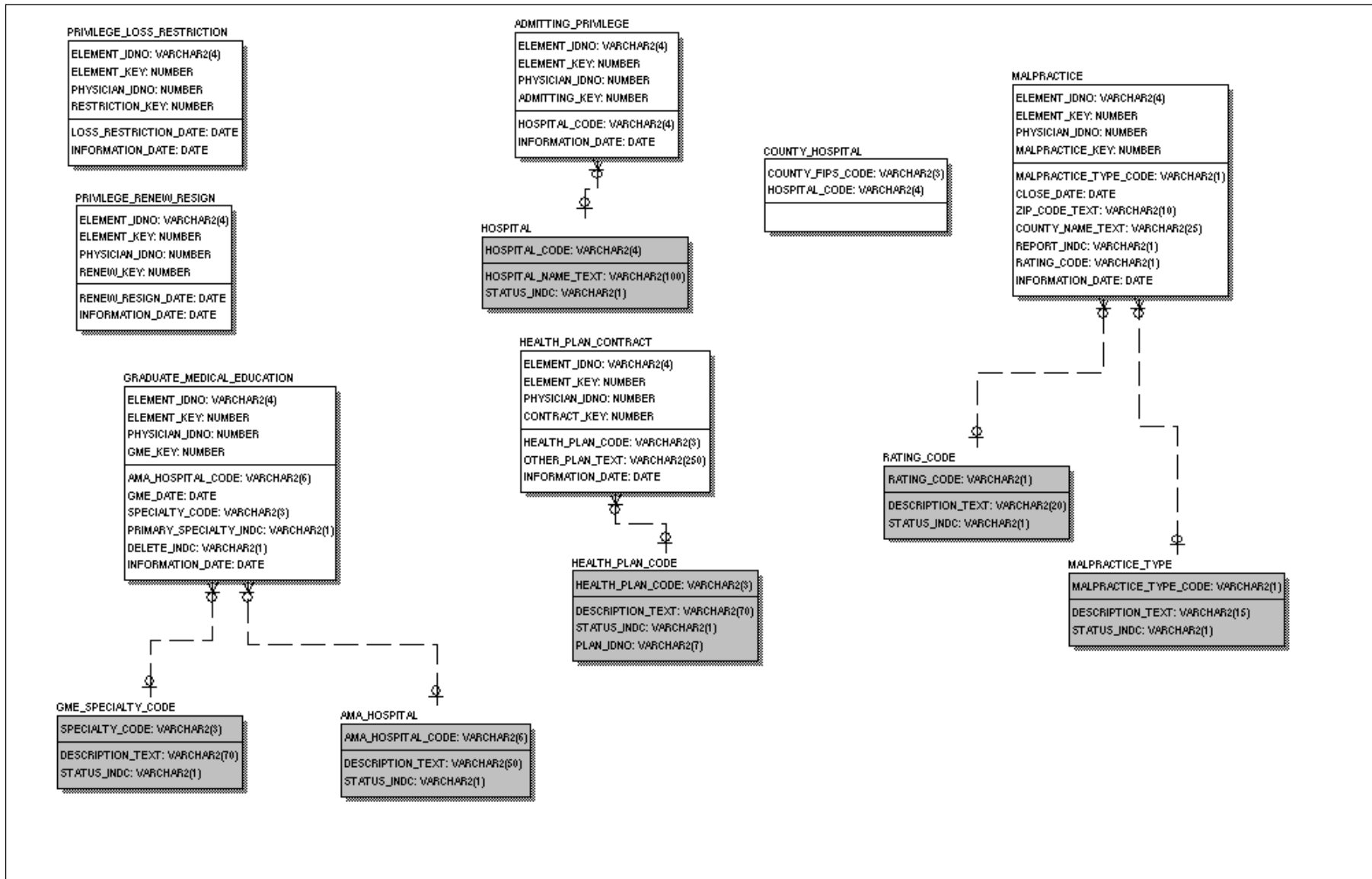
MALPRACTICE_STAGING

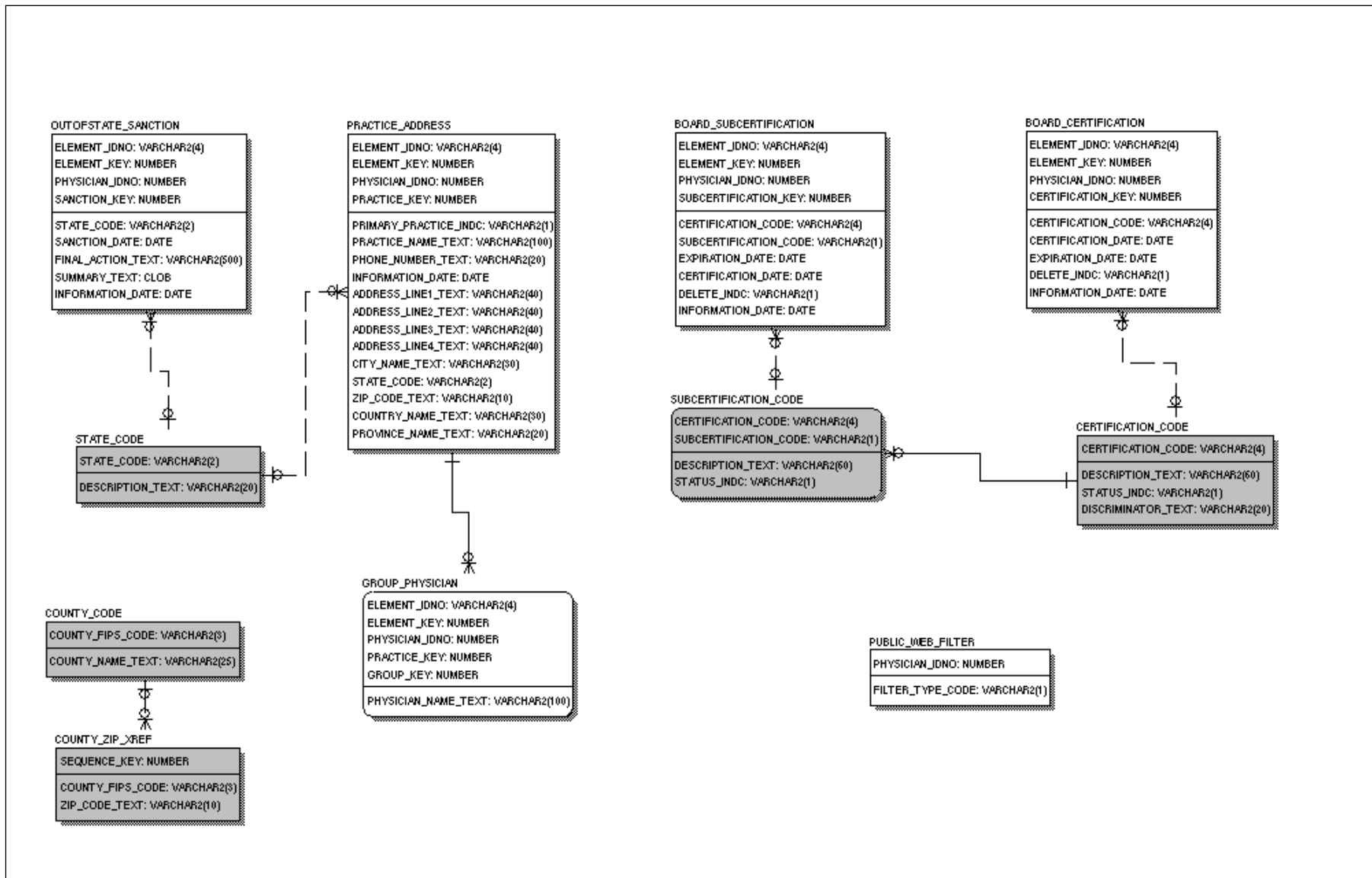
ELEMENT_IDNO: VARCHAR2(4)
ELEMENT_KEY: NUMBER
PHYSICIAN_IDNO: NUMBER
MALPRACTICE_KEY: NUMBER

MALPRACTICE_TYPE_CODE: VARCHAR2(1)
CLOSE_DATE: DATE
ZIP_CODE_TEXT: VARCHAR2(10)
COUNTY_NAME_TEXT: VARCHAR2(25)
REPORT_INDC: VARCHAR2(1)
RATING_CODE: VARCHAR2(1)
INFORMATION_DATE: DATE

Public Web Database







2. Data Dictionary

In this section, the relationships of the data attributes and tables to each other and to external sources and documents are detailed. Specifically, a detailed spreadsheet (NYPP Data Dictionary) is provided with the following information for the central repository. A second spreadsheet (Public Web Data Dictionary) is used to describe the public Web Database. Data definitions are only in the central repository spreadsheet as the public web is the current data subset of the central repository.

Entity or Table. Entity and table are synonymous terms that describe a related grouping of data elements that are referred to as attributes or columns. For example, a Name table may be comprised of last name, first name, and middle name columns as well as columns related to the source of the data and the identification of the doctor, by license number. For whom the specific information pertains. The following are the sections of the spreadsheet which relate to entity or tables:

- o **Entity Name.** The entity name is the plain text name of the table or grouping of elements.
- o **Table Name.** The system name for the entity or grouping of elements is the table name.
- o **Entity Definition.** This is a textual description of the table or entity.

Attribute or Column. Attribute and column are synonymous and are specific data elements. Attributes or columns can be filled through information provided from external data files, by information supplied by physicians, or can be generated by the Physician Profile system. Many column or attributes are used to relate other items or for systems actions and are never know by or visible to either the physician or public user. The following sections of the spreadsheet refer to attributes or entities:

- o **Attribute Name.** The attribute name is the text name of an individual element or column.
- o **Column Name.** The column name is the name used by the system for a single data element.
- o **Column Datatype.** The column datatype describes, in systems terms, the characters that comprise the column such as numbers or variable characters and specifies the maximum number of characters in the column or data element.
- o **Column Definition.** The column definition is a textual description of the column or data element.
- o **On Initial Profile Survey.** Attributes indicated with an “X” in this section will be contained on the pre-printed Profile Surveys sent to Physicians. Where “blank” is indicated, no information will be provided on the pre-printed Profile Surveys but space

will be provided for the physician to report the information. The physician web site will handle data similarly in that indicated attributes will be displayed and those with “blank” will allow for entry by the physician.

- 0 **On Profile Survey Review Copy.** Indicated attributes, those with an “X” under this section header, will be contained on the Profile Survey review copies sent to Physicians after receipt and data entry of their Profile Survey response. The physician web site will handle data similarly in that indicated attributes will be displayed after entry by the physician.
- 0 **On Profile Survey Codes Sheet.** Indicated attributes are either codes or the text descriptions of codes. If indicated, these attributes will be provided on the Profile Survey Codes Sheet so that the Physician can enter the code onto the Profile Survey. The physician web site will handle data similarly in that physicians will be provided a list of indicated text attributes to choose from, although the codes are not needed on the web and will therefore not be provided.
- 0 **On Public Web Site.** Indicated attributes will be viewable on the public web site.
- 0 **On Public Web Site if in Dispute.** The data for some attributes is provided by outside sources and physicians are not provided the opportunity to change the data. In some cases, a physician may disagree with and dispute the data provided. If this occurs, the disputed information will be sent to the data source and the data may or may not be presented on the public web site, as detailed in this section.

Initial Sources. The sections of the spreadsheet specify how the data will be incorporated into the database during the initial phases of the program.

- 0 **Self Report Only.** All indicated attributes will only reflect self report information from physicians.
- 0 **Pre-populate Source.** This information indicates the primary source for pre-populating data for Profile Surveys and the physician web site.
- 0 **Alternate Pre-populate Source.** This information indicates the secondary source for pre-populating data for Profile Surveys and the physician web site. Secondary source data will be used only when primary data is not provided unless otherwise specified in Notes.
- 0 **Doctor can Overwrite/Dispute.** Pre-populated data can either be changed or disputed by physicians. Physicians will be allowed to change, on either hardcopy or web Profile Surveys, pre-populated data for attributes indicated as “Overwrite”. Conversely, physicians will not be allowed to change pre-populated attributes that indicate “Dispute”.
- 0 **Call Center Counselor Enters.** These data elements can be input by call center or helpline counselors.

- ⊖ **Disputed Info Sent to.** This section provides the data provider that is notified if a physician disputes their data for the attribute.

Updates. These sections detail actions related to on-going changes after initial verification in the early phases of the program.

- ⊖ **External Sources.** This section indicates which attributes will be updated by periodic files received from various data sources and what the source is for the update of the specific attribute.
- ⊖ **Can be Updated on Website Without Doctor’s Review.** Indicated attributes can be displayed on the public web site based on updates received from indicated sources without verification from physicians. Conversely, attributes on the public web site that are not indicated here will not be updated on the public web site until physician verification is obtained.
- ⊖ **Overwrite Doctor Provided with Other Source.** This section specifies which attribute’s data, if provided by doctors, can be replaced by data from external sources and under what conditions a replacement can occur.
- ⊖ **Doctor can Change/Dispute.** Physicians will be allowed to change on an on-going basis, on either hardcopy or web Profile Surveys, data for attributes indicated as “Change”. Conversely, physicians will not be allowed to change attributes that indicate “Dispute” but instead may dispute the data.
- ⊖ **Notes.** These provide further information on the attribute or entity.

3. External Files and Import Procedures

The key external source for the Physician Profile database is the State Education Department (SED) file. The SED file is described in the following spreadsheets (Data File Exchange Formats, SED weekly and SED quarterly). Specifically, physician records are created based on data in the SED files. The Physician Profiling Database does not retrieve information from other sources unless there is already a record of the physician, derived from the SED file, in the database.

Information from SED records is loaded into the database if $M\text{-OFF} = 1$ and $HD\text{-STATUS} = 1$. Specifically, a new record is created if the license number is not found in the Physician Profile database. If a record already exists for the license number, changed data is input into the records, subject to the rules in Data Dictionary, if applicable.

Address fields are loaded from SED records if $M\text{-OFF} > 1$ and the license number in the record already exists in the Physician Profile database. These addresses, from SED records when $M\text{-OFF} > 1$ are for research purposes only and will not be displayed.

Data from all other sources is captured by the Physician Profile database only if a record with the same license number already exists. Those files are the AMA Physician Master file extract, the AOA Osteopath file extract, the OPMC Final Actions files, and the Medical Malpractice (MedMal) file from OPMC. If a record already exists in the Physician Profile database for the license number, changed data is input into the records, subject to the rules contained in the spreadsheet (Data Dictionary) noted above. File layouts and definitions for AMA and AOA may be obtained from AMA and AOA.