The State of Florida's Department of Education utilizes the following methods for creating unique identifiers for the purpose of data matching. Information is shared with the state's permission.

RCT-011 Identifying a student with demographic information

This rule identifies and links a unique data warehouse student identifier for each student record containing demographic data from various data sources such as prek-12, community colleges, state universities and others. As a result of this unique identification, all information related to an individual student such as courses, grades, financial aid and so on is linked to the student longitudinally across the various educational institutions the student may attend over a period of time.

Student identification is done by comparing demographic data elements to a "Student Identification Reference" table which contains a unique student identifier used internally for the data warehouse. The demographic data elements, from PreK-12, Community Colleges, State Universities, etc., used to identify a student are as follows:

Input File Format for Student Identification (Records to be linked)			
Data element	**Description**	**Format**	**List of values**
Data source	Identifies the sources of the data.	X(04)	PK12 CC SUS…
Student ssn id	May contain the social security number used to identify a student.	9(9)	
Student non-ssn id	May contain a non-ssn student identification.	X(18)	Concatenation of the following data elements: • Data source x(04), • Institution code x(04), • Student identification number assigned.
Student alias ssn id	May contain the social security number used to identify a student.	9(9)	
Student alias non-ssn id	May contain a non-ssn student identification composed of the concatenation of data	X(18)	Concatenation of the following data elements: • Data source

Input File Format for Student Identification (Records to be linked)			
Data element	**Description**	**Format**	**List of values**
	source, institution code and the student identification number assigned.		x(04), • Institution code x(04), Student identification number assigned.
Last name	Student's last name	X(30)	
First name	Student's first name	X(30)	
Middle name	Student's middle name	X(30)	
Middle initial	Student's middle initial	X(01)	
Birth date	Student's birth date	Date	
Gender	Student's gender	X(01)	
Racial category	Student's racial category	X(01)	

The "Student Identification Reference" table contains all the identifying data elements required to be associated with a unique student identifier that is assigned to a student for the data warehouse. The data elements stored in this table are described as follows:

Student Identification Reference Table Format			
Data Element	**Description**	**Format**	**List of Values**
Unique student identifier	Identifies the sources of the data.	9(12)	
Student ssn id	Contains the social security number used to identify a student.	9(9)	May come from student identification number or student identification number – alias
Student non-ssn id	Contains a non-ssn student identification.	X(18)	May come from student identification number or student identification number – alias
Last name	Student's last name	X(30)	
First name	Student's first name	X(30)	
Middle name	Student's middle name	X(30)	
Middle initial	Student's middle initial	X(01)	
Clean last name	Student's clean last name	X(30)	
Clean first name	Student's clean first name	X(30)	
Clean middle name	Student's clean middle name	X(30)	
Phonetic last	Student's phonetic last	X(04)	

Student Identification Reference Table Format			
Data Element	**Description**	**Format**	**List of Values**
name	name		
Phonetic first name	Student's phonetic first name	X(04)	
Phonetic middle name	Student's phonetic middle name	X(04)	
Birth date	Student's birth date	Date	
Gender	Student's gender	X(01)	
Racial category	Student's racial category	X(01)	

As well, a parameter-driven table is required to determine what rules are applied to the student identification process for the source data in question. The rules are:

- Through which rule should the searching for a unique student identifier be performed (Rules 1 through 8)?
- What is the access level to the "Student Identification Reference" table?
  - Update mode allows the creation of new unique student identifiers when required.
  - Query mode allows access to the reference table in read-mode only and, therefore, is not allowed to create new unique student identifiers.

An output file is generated as a result of processing the input file by associating a unique student identifier with each input record. The output file format is as follows:

Output File Format for Student Identification (Linked Records)			
Data Element	**Description**	**Format**	**List of Values**
Unique student identifier	Internal unique student identifier assigned to the student for the data warehouse	9(12)	
Identifier type	Determines whether the unique student identifier assigned to this record is newly created or a previously existing number.	X(1)	N – newly created unique student identifier E – previously existing unique student identifier
Data source	Identifies the sources of the data.	X(04)	PK12 CC SUS…
Student ssn id	Contains the social security number used to identify a student.	9(9)	
Student non-ssn	Contains a non-ssn student	X(18)	Concatenation of the

Output File Format for Student Identification (Linked Records)			
Data Element	**Description**	**Format**	**List of Values**
id	identification.		following data elements: • Data source x(04), • Institution code x(04), • Student identification number assigned.
Student alias ssn id	Contains the social security number used to identify a student.	9(9)	
Student alias non-ssn id	Contains a non-ssn student identification composed of the concatenation of data source, institution code and the student identification number assigned.	X(18)	Concatenation of the following data elements: • Data source x(04), • Institution code x(04), Student identification number assigned.
Last name	Student's last name	X(30)	
First name	Student's first name	X(30)	
Middle name	Student's middle name	X(30)	
Middle initial	Student's middle initial	X(01)	
Birth date	Student's birth date	Date	
Gender	Student's gender	X(01)	
Racial category	Student's racial category	X(01)	

Student identification:

The following name-matching rules are applied in the order they are presented in the following pages. Whenever a data element is not present for the search in question, the next hierarchical rule in the list is applied. The search rules are as follows:

1. Search by student non-ssn id or student alias non-ssn id.

2. Search by student ssn id or student alias ssn id and birth date.

3. Search by student ssn id or student alias ssn id, last name and first name.

4. Search by student ssn id or student alias ssn id, clean last name and clean first name.

5. Search by student ssn id or student alias ssn id, phonetic last name and phonetic first name.

6. Search by last name, first name and birth date.

7. Search by clean last name, clean first name and birth date.

8. Search by phonetic last name, phonetic first name and birth date.

There are three possible outcomes to any of the above-mentioned searches. They are as follows:

- A **No match found** result occurs when a search returns no unique student identifier from the "Student Identification Reference" table.

- A **Match found with one unique identifier** result occurs when a search returns one unique identifier from the "Student Identification Reference" table.

- A **Match found with several unique identifiers** result occurs when a search returns more than one unique identifier from the "Student Identification Reference" table.

No match found

- Execute the next hierarchical search rule until a match is found or all searches are performed.

- When all searches are performed and there is still no match then:
  - if the request is in **query mode**, the student record is rejected and a non unique identifier is linked to the record;
  - if the request is in **update mode**, a new unique student identifier is created in the "Student Identification Reference" table using the student's demographic data to create the occurrence, and the student record is linked to this new unique student identifier.

Match with one unique identifier

- The unique student identifier found is assigned to the student record in question.

- If any identifying data elements in the student record differ from those in the "Student Identification Reference" table, a new occurrence is created in the "Student Identification Reference" table using the same unique student identifier found.

Match with several unique identifiers (N match)

- The search is refined by using the remaining identifying data elements to search for a unique student identifier match. When all refined searches are exhausted and a single unique student identifier is not found, the most recently created unique identifier is linked to the student record.

- If one unique identifier is found after refining the search, the unique identifier found is linked to the student record in question.

- If any of the identifying data elements in the student record differ from those in the "Student Identification Reference" table, a new occurrence is created in the "Student Identification Reference" table using the same unique student identifier found.