

---

# CAAL-Skills

---

## Data Exchange Specification Requirements

Pilot Project Documentation



## Change Log

Version	Date	Author	Change
<b>1.00</b>	8/21/17	E. Holm P. Banas	CWDA Version
<b>1.10</b>	9/26/17	E. Holm P. Banas	<p>Updated based on review with CWDA teams:</p> <ul style="list-style-type: none"> <li>Revised DED #002 Enrollment Organization Name to reference the County Name</li> <li>Revised DED #003 Enrollment Organization ID to reference the County Code</li> <li>Revised DED #015 Participant Unique Individual Identifier to reference the CIN or Client Identification Number</li> <li>Revised DED #017 Participant Ethnicity/Race to Participant Ethnicity VCHAR(100) changed definition and field name</li> <li>Added DED #017-1 Participant Race</li> <li>Revised DED #020 to "Homeless Individual" (removing "Runaway Youth designation</li> <li>Removed DED #024 Participant English Language Learner</li> <li>Revised DED #027 Temporary / Permanent Disability to include County Code crosswalk</li> <li>Revised DED #029 Single Parent to include expanded description</li> <li>Removed DED #030 Long Term Unemployed</li> <li>Revised DED 035, 036, 064 to include CalFRESH and CalWORKs activities and services</li> </ul>
<b>1.20</b>	4/17/18	P. Getz	Added FY 17/18 to the Timeframe Section.

## Table of Contents

1. Overview.....	7
1.1. Project.....	7
1.2. Objectives .....	8
1.3. Metrics.....	9
1.3.1. Aggregation and Reporting Environment.....	9
1.3.2. Cross System Tracking .....	9
1.3.3. Program History.....	9
1.3.4. Employer Needs and Economic Trends.....	9
1.3.5. Commonality, Consistency, Compliance .....	9
1.3.6. Summary.....	10
1.4. Scope .....	10
1.4.1. Partners .....	10
1.4.2. Timeframe .....	10
1.4.3. Core Entities and Attributes .....	11
2. Data Diagrams .....	13
2.1. Conceptual Data Model.....	13
3. Data Elements .....	14
3.1. Data Types .....	14
3.1.1. Maximum Lengths .....	14
3.1.2. Null Values .....	15
3.1.3. Currency .....	15
3.1.4. Case .....	15
3.1.5. Date Format.....	15
3.1.6. Binary Values .....	15
3.1.7. Unknown/Not Applicable/Other Values .....	15
3.1.8. Field Formatting .....	15
3.1.9. Reference Codes.....	16
3.2. Identifiers and Definitions.....	16
3.3. Data Element Name.....	16
3.4. Date Range .....	16

3.5.	Data Relationships .....	16
3.5.1.	Primary and Foreign Keys .....	16
3.5.1.1.	Participant Example .....	17
3.6.	Participant Individual Record Layout (PIRL) .....	17
4.	Data Extraction .....	18
4.1.	File Format .....	18
4.2.	File Specification .....	18
4.2.1.	Character Encoding .....	18
4.2.2.	Character Case .....	18
4.2.3.	Implementation of Delimited Separated Values .....	19
4.2.4.	File Encryption .....	20
4.2.5.	File Organization .....	20
4.2.6.	File Sizing .....	21
4.2.7.	File Naming and Sequencing .....	21
4.2.8.	Folder Structure .....	23
4.3.	Data Transfer .....	24
4.3.1.	Introduction .....	24
4.3.2.	Device .....	24
4.3.3.	Advanced Encryption Standard (AES) .....	25
4.3.4.	Federal Information Processing Standards Certification .....	26
4.3.5.	Hardware Encryption .....	26
4.3.6.	Device Format .....	27
4.3.7.	Device Identification .....	29
4.3.8.	Device Loading .....	29
4.3.9.	Device Chain of Custody .....	29
5.	Other Project Agreements .....	32
5.1.	Data Sharing Agreements .....	32
6.	Partner Response .....	33
6.1.	High Level Tasks and Key Dates .....	33
6.2.	High Level Tasks and Costs .....	33
6.3.	Response Timeframe .....	33
7.	Contact Information .....	34
7.1.	CAAL-Skills Project – Chief Technical Architect .....	34

7.2.	Primary Technical Contact.....	34
7.3.	Alternate Technical Contact .....	34
8.	Data Element Dictionary.....	35
8.1.	Introduction.....	35
	Partner Identification .....	36
	Partner ID .....	36
	Enrolling Organization .....	37
	Enrollment Organization Name .....	37
	Enrollment Organization ID .....	38
	Funding Source .....	39
	Fund Source .....	39
	Participant Identification .....	40
	Social Security Number .....	40
	Participant First Name .....	41
	Participant Middle Name .....	42
	Participant Last Name .....	43
	Participant City .....	44
	Participant Postal Code .....	45
	Participant State Code .....	46
	Participant Date of Birth.....	47
	Participant Sex / Gender.....	48
	Participant Unique Individual Identifier .....	49
	Participant Demographics .....	50
	Participant Ethnicity / Race .....	50
	Participant Veteran Status.....	52
	Homeless Individual / Runaway Youth.....	53
	Participant Foster Care Youth .....	54
	Participant English Language Learner .....	<b>Error! Bookmark not defined.</b>
	Temporary / Permanent Disability .....	55
	Migrant / Seasonal Farmworker.....	56
	Participant Single Parent .....	57
	Long-Term Unemployed.....	<b>Error! Bookmark not defined.</b>
	Workforce Training and Related Education .....	58

Type of Career Service .....	58
Employment and Training Service Related to Supplemental Nutritional Assistance Program (SNAP E&T) or Welfare to Work .....	60
Type of Training Service .....	61
Supportive Services .....	63
Type of Supportive Service or Activity .....	63
First Date of Service or Activity .....	65
Last Date of Service or Activity .....	66

## 1. Overview

### 1.1. Project

Federal and state law required the development of a “State Plan” to address workforce development, training and related challenges. Under the direction of the California Labor and Workforce Development Agency (LWDA) the plan was developed by the California Workforce Development Board (CWDB).

Implementation of the State Plan includes the development of a comprehensive, integrated, interoperable data system that measures and assesses program outcomes in a systematic and efficient manner. The data system’s primary objective will be helping the CWDB and State Plan Partners meet statutorily mandated performance reporting requirements while also developing the State’s capacity to evaluate and assess program success.

This will enable the CWDB and its partners to develop policies which facilitate the movement of people into jobs “providing economic security or job placement in an entry-level job that has a well-articulated career pathway or career ladder to a job providing economic security.”[1] State law defines these jobs as those that provide, “a wage sufficient to support a family adequately, and, over time, to save for emergency expenses and adequate retirement income, based on factors such as household size, the cost of living in the worker’s community, and other factors that may vary by region.”

Workforce development, training, and related education programs provide job-specific training, basic skills education, and related support services to help individuals participate in civic life and the labor market. California spends over \$6 Billion annually on workforce development, training, and related education programs. Over three million people access one or more of California’s workforce development, training, and related education programs annually.

These programs and the data relevant to them are administered by the following organizations:

1. Labor & Workforce Agency
  - a. Employment Development Department – Workforce Services Division
  - b. Employment Development Department – Labor Market Information Division
  - c. Employment Development Department – Tax Branch
  - d. Employment Training Panel
  - e. Department of Industrial Relations, Division of Apprenticeship Standards
  - f. Workforce Development Board
2. Public School System - Department of Education, Career and College Transition Division (includes Career Technical Education and Adult Education)
3. Community College Chancellor’s Office - Technology, Research and Information Systems (TRIS) Division
4. Health & Human Services Agency
  - a. Department of Social Services – CalFresh
  - b. Department of Social Services - CalWORKs
  - c. Department of Rehabilitation
5. County Welfare Directors Association of California

Workforce development, training, and related education providers are required to report information about their program, program participants, and the participant outcomes. Currently, the outcome

measurements and data reported varies from program to program though efforts are underway to align program performance measures. Historically, data collection has been used almost exclusively for oversight and compliance monitoring, yet program and outcome information is essential for effective decision-making. For example:

- State leadership, including the Governor and Legislature, require workforce training and education information that will enable them to develop evidence-based statutes and regulations to facilitate the effective allocation of scarce resources.
- Policy Makers require information regarding significant barriers to employment, such as a significant disability, that may require intensive job preparation.
- Workforce development, training, and related education administrators require information enabling them to evaluate program effectiveness, shape policies, and develop linkages among institutions.
- Teachers and guidance counselors require information to develop curricula, assess and guide students, and create learning plans.
- Workforce professionals require information to identify the remedial needs of students, provide market-informed career counseling, and help clients find jobs.
- Students, parents, and job seekers require information to make informed education, training, and career decisions.

The purpose of the CAAL-Skills project is to develop a comprehensive, integrated, interoperable data system that measures and assesses program outcomes in a systematic and efficient manner. The data system's overarching purpose will be helping the CWDB (and its State Plan partners) meet statutorily mandated performance reporting requirements and align with Workforce Innovation and Opportunity Act (WIOA) of 2014 and AB 1336 requirements. WIOA is a publicly-funded workforce system that is a national network of organizations which provide a range of employment, education, training, and related services and supports - to help jobseekers secure good jobs while providing businesses with the skilled workers they need to compete in the global economy.

The CAAL-Skills project objectives include providing automated functionality that will improve the State's capacity to evaluate and assess Workforce Development programs' success. This will enable the CWDB and its partners to develop policies which facilitate the movement of people into jobs providing economic security or job placement in jobs with a well-articulated career pathway or career ladder to a job providing economic security.

## 1.2. Objectives

The information required to facilitate effective decision-making as specified in the project overview must:

- Include complete program and participant information;



- Facilitate the tracking of individuals across institutions, through programs and into the labor market;
- Consistently evaluate program measurements and other information and variables across time and between programs;
- Be updated on a regular basis;
- Allow for the analysis of individual and aggregated information;
- Be accessible to State leadership; state and local workforce planning, education administrators and others; and,
- Be secure - the solution, including any data sharing, will comply with applicable state and federal privacy and data security requirements and best practices.

### 1.3. Metrics

A repository of Cross-System Analytics and Assessment for Learning Skills attainment (CAAL-Skills) will be created in a pilot environment to meet the above objectives. The primary metrics of the pilot will be the following:

#### 1.3.1. Aggregation and Reporting Environment

The target reporting environment includes the collection, consolidation and reporting of workforce development, training and related education information from key workforce data partners.

#### 1.3.2. Cross System Tracking

A method will be developed to link individuals across programs even when one or more of the source systems does not include an individual's social security number.

#### 1.3.3. Program History

The targeted environment includes consolidated reporting of an individual's workforce development, training and related education program involvement history..

#### 1.3.4. Employer Needs and Economic Trends

The target environment includes reporting on regional employer needs in comparison to current and planned workforce development, training and related education offerings.

#### 1.3.5. Commonality, Consistency, Compliance

The target system will:

- a. Apply common measures and consistent data-element definitions across programs;
- b. Apply a unique participant identifier that can be used to link an individual appearing in two or more systems;
- c. Invoke a consistent understanding and approach for protecting the personally identifiable information (PII) of program participants;
- d. Simplify the exchange of data; and,
- e. Comply with federal laws, state laws, state regulations and best industry practices for information security and privacy.

#### 1.3.6. Summary

The CAAL-Skills Pilot project will create a central repository to host data collected by the project Data Sharing Partners. The data will be utilized to generate information and metrics required by California Unemployment Insurance Code Section 14013, as well as provide a mineable warehouse for further analytics.

### 1.4. Scope

The scope of the CAAL-Skills Pilot Project includes receiving data from state and local workforce partners as defined in Data Sharing Agreements.

#### 1.4.1. Partners

This project involves the agencies from across the Workforce services spectrum including education, training, and supportive services. The breadth of participation ensures the data necessary to meet the project requirements and provide a rich data repository. The Data Sharing Partners for the pilot portion of the CAAL-Skills project include:

- California Community Colleges Chancellor's Office (CCCCO)
- California Department of Education (CDE)
- California Department of Social Services (CDSS)
- California Department of Industrial Relations (DIR)
- California Department of Rehabilitation (DOR)
- California Employment Development Department – Central Office Workforce Services Division (EDD-COWSD)
- California Employment Development Department – Labor Market Information Division (EDD-LMID)
- California Employment Development Department – Tax Branch (EDD-TAX)
- Employment Training Panel (ETP)

#### 1.4.2. Timeframe

The fiscal years 2013 – 2014 through 2017 – 2018 will be used in the baseline. For clarity, the fiscal years include the following data ranges:

Fiscal Year 2013-2014 (FY13-14) = July 1, 2013 through June 30, 2014

Fiscal Year 2014-2015 (FY14-15) = July 1, 2014 through June 30, 2015

Fiscal Year 2015-2016 (FY15-16) = July 1, 2015 through June 30, 2016

Fiscal Year 2016-2017 (FY16-17) = July 1, 2016 through June 30, 2017

Fiscal Year 2017-2018 (FY17-18) = July 1, 2017 through June 30, 2018

If data elements are not available within that timeframe, the partner will be asked to submit the available elements. A method will be developed to link individuals across programs even when one or more of the source systems does not include an individual's social security number.

#### 1.4.3. Core Entities and Attributes

The core entities and attributes present an overview of the data to be gathered and modeled. The higher level in this hierarchy are the core entities – groupings of attributes and/or distinctly separate objects. The lower level in this hierarchy are the attributes – or characteristics of the core entities. *Note: These are concepts not concrete definitions. Some core entities and attributes may not apply to specific Data Sharing Partners.*

- Fund Source – Funds used to pay for participant services (e.g. WIOA Title 1 Adult)
- Partner – Data Sharing Partner and/or Administrative Agency (e.g. EDD-COWSD)
- Enrollment Organization – Organization who determines participant eligibility and/or enrolls in service. May be the same as the Partner. (e.g. Local Workforce Board)
- Provider – Organization who offers services to participant. May be the same as the Enrollment Organization or may be contracted by Enrollment Organization. (e.g. Onestop)
- Participant
  - Identification
  - Demographic
  - Employment Barriers
- Services
  - Education and Training Services – Occupational or general education (e.g. Adult Basic Education, Career Technical Education)
  - Career Services – Employment related services ( e.g. Resume writing workshop, career counseling)
  - Supportive Services – Services to support participation in Education, Training, and Career Services (e.g. CalFresh or CalWORKS benefits)
- Service Outcomes
  - Certificate/Skill/Degree Attained
  - Start Date
  - Completion/Withdraw Date

- Employment Outcomes
  - Employment Attained
  - Wages (Participant income from Tax branch records)
    - 2<sup>nd</sup> Quarter
    - 4<sup>th</sup> Quarter

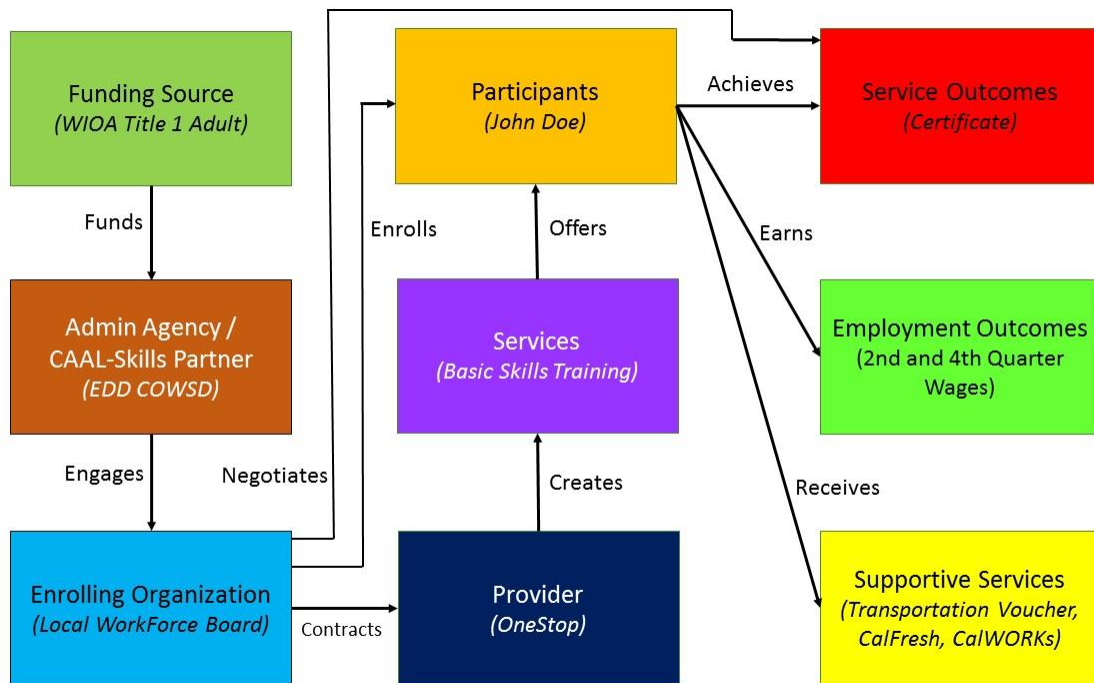
Further details regarding the exact data elements included as attributes and their relationships will be provided in the specification section of this document.

## 2. Data Diagrams

### 2.1. Conceptual Data Model

A conceptual data model is a map of concepts and their relationships used for databases. This describes the semantics of an organization and represents a series of assertions about its nature. Specifically, it describes the things of significance to an organization (entity classes), about which it is inclined to collect information, and characteristics of (attributes) and associations between pairs of those things of significance (relationships).

Because a conceptual schema represents the semantics of an organization, and not a database design, it exists on various levels of abstraction. The conceptual data model for the CAAL-Skills project is below. It is offered to assist in an overall understanding of the types of data to be collected and their relationships.



Each of the above entities (blocks) in the conceptual model may be represented in the data extracted from the CAAL-Skills partner's source systems and transferred to the CAAL-Skills Pilot project team for integration into the central workforce and outcomes repository. Not all partners will offer all entities or data elements, just those related to their services and/or benefits.

## 3. Data Elements

### 3.1. Data Types

A data type or simply “type” is a classification of data which tells the database, application, compiler, interpreter, or alternative data store how the programmer or administrator intends to use the data. Modern programming languages and data stores support various types of data (i.e. real, integer, or Boolean). A data type provides a set of values from which an expression (i.e. variable, function) may take its values. The type defines the operations that can be done on the data, the meaning of the data, and the way values of that type can be stored.

The CAAL-Skills Pilot Project will utilize a limited subset of available data types to reduce complexity and simplify the formatting process for moving data from the source system to the secure data transfer device. Below are the four basic data types and their formats:

- Variable Character (VCHAR) – For storing variable length alpha-numeric data. (e.g. Last Name)
- Integer (INT) – For storing numeric values not requiring fixed precision or scale. (e.g. Postal Code)
- Numeric (NUM) – For storing numeric values requiring fixed precision and scale. (e.g. \$1,432.10)  
*Note: Precision is the number of digits in a number. Scale is the number of digits to the right of the decimal point in a number. For example, the number 123.45 has a precision of 5 and a scale of 2.*
- Date (DATE) – For storing month, day, and year. (e.g. January 18, 2015)

#### 3.1.1. Maximum Lengths

The maximum length of a given data element will be specified within the data element dictionary section for VARCHAR, INT, NUM data types. All elements will be considered variable length and will not require left or right padding utilizing spaces, special characters, or zero-filled positions. Data which exceed the maximum length should be truncated to meet the requirements of that particular data element. An example of each data type is listed below:

- Variable Character (VCHAR)
  - Data Element: Last Name
  - Maximum Length: 40 characters
  - Format: VCHAR (40)
  - Examples: SMITH, JOHNSON-KEENER
- Integer (INT)
  - Data Element: Postal Code
  - Maximum Length: 9 positions
  - Format: INT
  - Examples: 94518, 10027, 85602-1023

- Numeric (NUM)
  - Data Element: 2<sup>nd</sup> Quarter Earnings
  - Maximum Length: 9 precision, 2 scale
  - Format: NUM (9,2)
  - Examples: \$1437.99 , \$7,210.33

### 3.1.2. Null Values

Data elements included in the specification for which values do not exist in the source system or cannot be translated using the data element cross reference should be left blank as noted in the CSV file format of: "DataElement1" || "DataElement3". Do not include spaces, double quotes without enclosed values, zeroes, or other special characters. The presence of the field separators without data will be sufficient for the target system in CAAL-Skills to determine the lack of a value for a given data element.

### 3.1.3. Currency

Currency values should be treated as numeric (NUM) with precision and scale. Do not include currency symbols in the field.

### 3.1.4. Case

All alphabetic characters should be rendered in the upper or majuscule case. Do not provide lower or mixed case in VCHAR elements.

### 3.1.5. Date Format

The DATE data type should include only month, day, and year. It should not include timestamps. The format for DATE data types is: MMDDYYYY. The only exception to zero filling/padding of fields occurs in this format. Single digit months (January – September) should be left zero padded as should single digit days (1-9).

### 3.1.6. Binary Values

The data specification requirement defines binary values such as Veteran Status (Vet/Not a Vet). A value of 1 will be used to signify a positive response and a value of 0 will be used to signify a negative response. To provide the most flexibility for future program data, these data elements will utilize an INT data type.

### 3.1.7. Unknown/Not Applicable/Other Values

For "Unknown" values will utilize "99" to clearly note that response. For "Not Applicable" values will utilize "98". "Other" values, that is values outside of the list provided, will utilize "97". Note: These values are appropriate for INT and VCHAR data types.

### 3.1.8. Field Formatting

Formatting of special use numeric and dates is not to be performed. Social Security Numbers should just utilize the integers not the typical format: XXX-XX-XXXX. Phone numbers should not include spaces, parentheses, and dashes: (XXX) XXX-XXXX. Dates should not include positional

separators: MM-DD-YYYY. Unless specified in the Data Extraction Requirements all fields should contain the numeric or alpha characters only.

#### 3.1.9. Reference Codes

Due to the multiple source data systems for the CAAL-Skills repository, we will request reference or translation codes for some data elements to standardize responses and quickly assimilate source data into the repository. Those reference codes will be noted in each element. In some cases, multiple source data elements will map to a single CAAL-Skills reference code.

CAAL Skills values (as noted in the data dictionary) will be assigned to each data element. These values are recommended be used and collected for the CAAL Skills data repository. By establishing universal values, and cleansing the data to record those values, the data can be linked amongst the programs.

The values recommended identify the available options for the data element, including if the data element is not applicable or unknown to the Partner. The recommended values take into account the current values used by the Partners and provides for the least amount of data cleansing.

### 3.2. Identifiers and Definitions

Each data element in the data dictionary section will include a CAAL-Skills data element unique identifier, descriptive name, and full description of the element. These are provided for discussion and reference purposes only. This information should not be included in the data provided to the CAAL-Skills Pilot Project Team.

### 3.3. Data Element Name

Each data element in the data dictionary section will include a unique, concatenated CAAL-Skills data element name which is to be used in the header row of each file to identify the data associated with that element. As noted in the section above describing the CSV file format, the header row should contain a list of all data elements included in the file, in the order of appearance on the row line, and separated by a pipe (|) character.

### 3.4. Date Range

The scope of the data requested from the Data Sharing Partners will be specific for each partner and each data element. Please refer to the matrix at the beginning of the Data Element Dictionary. If the data is unavailable for a portion of the requested time frames, please provide only the available time frames in the data files sent to the CAAL-Skills project team. Inserting blank rows for incomplete or missing calendars years is not required.

### 3.5. Data Relationships

#### 3.5.1. Primary and Foreign Keys

In order to capture and maintain the relationship between core entities, the Data Sharing Partners will be asked to include their internal identifier for data elements (noted as Primary or Foreign keys) to generate unique and consistent identifiers for those elements. The identifiers may be numeric or alphanumeric. We will also request the original identifier for some data elements, but those will be considered distinctly separate data elements.



#### 3.5.1.1. Participant Example

The example below assumes a participant, John Doe, with a unique source system identifier of K109-0563 was enrolled in Career Technical Education at Sacramento City College in the September 1, 2015. Further, the participant completed this education by the end of December 30, 2015.

##### Elements and Values

**Participant:** John Doe, ID #K109-0563 (Unique Participant ID\* from source system)

**Service:** Career Technical Education - 9/1/15 (Training Program Code # 123456)

**Location:** Sacramento Community College (Location # 489)

**Service Outcome:** Certificate Attained - 12/30/15

In the above example, the Data Sharing Partner could use the Source System ID within all three entities to identify John Doe.

**Participant Entity:** John Doe, K109-0563, 3835 Freeport Blvd, Sacramento, CA

**Services Entity:** K109-0563, RW013, 489, 9/01/2015

**Services Outcome Entity:** K109-0563, RW013, 489, 12/31/2015

*\*The unique Participant ID may be the original source system identifier or a unique identifier created during the extraction process and maintained throughout the course of the data file creation.*

#### 3.6. Participant Individual Record Layout (PIRL)

Within the Data Element dictionary there is a comparison of the CAAL Skills data element and the related Department of Labor's (DOL) data element, as identified in Version 25 of the Participant Record Layout (PIRL). This comparison includes:

- A. The related DOL PIRL data element:
  - i. Number
  - ii. Name
  - iii. Type and Field Length
  - iv. Definition / Instructions; and
  - v. Code Values; and
- B. The Comparative Analysis.

## 4. Data Extraction

### 4.1. File Format

The pilot project will leverage the lowest-common-denominator in file formats to receive data from the Partners. Formats that use delimiter-separated values (DSV) store two-dimensional arrays of data by separating the values in each row with specific delimiter characters. Most database and data store programs are able to read or save data in a delimited format. Due to their wide support, DSV files can be used in data exchange among many applications and systems.

A delimited text file is a text file used to store data, in which each line represents a single row or record, and each line has fields separated by the delimiter. Compared to the type of flat file that uses spaces to force every field to the same width, a delimited file has the advantage of allowing field values of any length.

Any character may be used to separate the values, but the most common delimiters are the comma, tab, and colon. The vertical bar (also referred to as pipe) and space may also be used. In a comma-separated values (CSV) file the data items are separated using commas as a delimiter, while in a tab-separated values (TSV) file, the data items are separated using tabs as a delimiter. Column headers are sometimes included as the first line, and each subsequent line is a row of data. The lines are separated by newlines or carriage returns (CR/LF)

This project will reference the Internet Engineering Task Force (IETF) standard identified by the Request for Comment document RFC 4180. A Request for Comments (RFC) is a formal document from the IETF that is the result of committee drafting and subsequent review by interested parties. Some RFCs are informational in nature. Of those that are intended to become Internet standards, the final version of the RFC becomes the standard and no further comments or changes are permitted. Change can occur, however, through subsequent RFCs that supersede or elaborate on all or parts of previous RFCs. The full specification is documented here: <https://tools.ietf.org/html/rfc4180>

This project will deviate slightly from the standards published in RFC 4180 by replacing the comma used as a field separator with a pipe ( | ) character. This modification will accommodate commas within the fields. The full specification for the file is listed in the next section.

### 4.2. File Specification

#### 4.2.1. Character Encoding

The character encoding scheme for the files to be exchanged will be UTF-8. It is a character encoding scheme capable of encoding all possible Unicode code points. The encoding is defined by the Unicode standard with the full standard listed here:

<http://www.unicode.org/versions/Unicode6.0.0>

The encoding is variable-length and uses 8-bit code units. It was designed for backward compatibility with ASCII and to avoid the complications of endianness and byte order marks in the alternative UTF-16 and UTF-32 encodings.

#### 4.2.2. Character Case

All alphabetic characters are to be generated in upper case also referred to as majuscule. Mixed case and lower case within the source data should be converted to upper case for the extract files.

#### 4.2.3. Implementation of Delimited Separated Values

The specifications for implementing delimited separated values are as follows:

1. Each field (data element) is to be separated by a pipe (|) character. Unicode UTF-8 standards refer to this character as the “vertical line”. It is officially the 0x7C hex character also represented as U+007c. For example:

```
aaa | bbb | ccc
```

*Note: Spaces between the data values and the delimiter are not necessary and should not be included – the spaces in the example above are for clarity’s sake only.*

2. Each record is located on a separate line, delimited by a line break (CRLF). For example:

```
aaa | bbb | ccc CRLF  
zzz | yyy | xxx CRLF
```

3. The last record in the file should not have an ending line break. For example:

```
aaa | bbb | ccc CRLF  
zzz | yyy | xxx
```

4. There should be a header line appearing as the first line of the file with the same format as normal record lines. This header will contain names corresponding to the fields in the data specification (column titled “CAAL-Skills Field”) and should contain the same number of fields as the records in the rest of the file. For example:

```
field_name1 | field_name2 | field_name3 CRLF  
aaa | bbb | ccc CRLF  
zzz | yyy | xxx CRLF
```

5. Within the header and each record, there may be one or more fields, separated by commas. Each line should contain the same number of fields throughout the file. Spaces are considered part of a field and should not be ignored. The last field in the record must not be followed by a comma. For example:

```
aaa | bbb | ccc
```

6. Each field should not be enclosed in double quotes unless the field contains the delimiter as valid data within that field. If fields are not enclosed with double quotes, then double quotes may not appear inside the fields. For example:

```
"aaa|" | "|bbb" | "c|cc" CRLF  
zzz | yyy | xxx
```

7. Fields containing line breaks (CRLF), double quotes, and pipes should be enclosed in double-quotes. For example:

```
"aa""a" | "b CRLF bb" | "c|cc" CRLF  
Zzz | yyY | xxx
```

8. If double-quotes are used to enclose fields, then a double-quote appearing inside a field must be escaped by preceding it with another double quote. For example:

```
aaa | "b""bb" | ccc
```

9. The Augmented Backus-Naur Form (ABNF) grammar appears as follows:

- file = [header CRLF] record \*(CRLF record) [CRLF]
- header = name \*(PIPE name)
- record = field \*( PIPE field)
- name = field
- field = (escaped / non-escaped)
- escaped = DQUOTE \*(TEXTDATA / PIPE / CR / LF / 2DQUOTE) DQUOTE
- non-escaped = \*TEXTDATA
- PIPE = 0x7C
- CR = %x0D

10. Security considerations: DSV files contain passive text data that should not pose any risks. However, it is possible in theory that malicious binary data may be included in order to exploit potential buffer overruns in the program processing DSV data. All DSV files received by the pilot program will undergo CRC, virus, malware, and malicious code detection and quarantine (if necessary) prior to extracting.

#### 4.2.4. File Encryption

The devices utilized to copy and/or extract the source data files at the partner facilities, host those files in transit, and load the file contents into the CAAL-Skills pilot project central repository will be encrypted with full-disk AES 256-bit hardware-based encryption. More details on the devices will be offered in a following section. Due to the use of hardware-based encryption the data files will not require additional encryption.

#### 4.2.5. File Organization

Each entity identified in the Conceptual Data Model will be represented by one or more files. The contents of each file will be the elements (fields) listed in the Data Element section. The entities are as follows:

1. Funding Source
2. Administrative Agency / Partner
3. Enrolling Organization
4. Provider

5. Services
6. Participants
7. Supportive Services
8. Services Outcomes
9. Employment Outcomes

Data Sharing Partners will generate only the files applicable to their organizations. The actual elements for each file will be determined by the “File” classification field for each data element. For example:

**Data Element:** Participant First Name  
**File:** Participants

**Data Element:** Provider Postal Code  
**File:** Providers

#### 4.2.6. File Sizing

The encrypted data transfer devices will have a 64 GB capacity which will be sufficient to host and transport the requested information. Maximum file sizes are indicated as a matter of ease of handling and importation by the CAAL-Skills pilot project data integration system. Data Sharing Partners are encouraged to utilize file sizes which accommodate their source systems and data extraction processes – that is to say, if smaller files are more easily generated and managed then the Data Sharing Partners may further reduce these recommended sizes as needed. Minimum file sizes are listed below to ensure files are concatenated to improve efficiency in the data sharing processes.

ID	Entity / File Classification	Maximum Size	Minimum Size
1	Funding Source	250MB	1KB
2	Admin Agency / Partner	250MB	1KB
3	Enrolling Organization	250MB	1KB
4	Provider	250MB	1KB
5	Services	250MB	1KB
6	Participants	250MB	1KB
7	Supportive Services	250MB	1KB
8	Services Outcomes	250MB	1KB
9	Employment Outcomes	250MB	1KB

#### 4.2.7. File Naming and Sequencing

The file naming and sequencing conventions for the pilot project will assist in identifying the files both at the source as they are created and placed on the encrypted transfer device and when they are received by the CAAL-Skills Pilot Project Team.

File Naming will identify the following characteristics of each file transmitted. Again, not every Data Sharing Partner will generate every file type listed. Within the Naming Format the etymology reference is: “X” = alpha character (A-Z only), ‘N’ = numeric (integer), and MMDDYYYY = date. A number (1-10) following the alpha character or “X” in the naming format indicates a fixed length. Multi-positional numeric or integers should be left zero filled (e.g. 001, 013). All file characteristics will be separated by a single underscore (\_). All file names should be in upper case or majuscule.

Position	File Characteristic	Naming Format	Default Value
1	CAAL-Skills Project	XXN	CS1
2	Date of File Creation	MMDDYYYY	N/A
2	Entity or File Classification*	XXXX	N/A
3	Sequence of Entity or File Classification (for multiple files)	NNN	001
4	CAAL-Skills Request ID	XNN	R01
5	File Extension	XXX	CSV
6	Data Sharing Partner – <i>this is not included as the data transfer device will indicate the partner</i>		
*	<i>For naming refer to the chart below</i>		

ID	Entity / File Classification Naming	File Characteristic
1	Funding Source	FUND
2	Administrative Agency / Partner	AAPA
3	Enrolling Organization	ENRO
4	Provider	PROV
5	Services (training and education)	SERV
6	Supportive Services	SUPP
7	Participants	PART
8	Service Outcomes	OUTC
9	Employment Outcomes	WAGE

Examples of file naming are listed below to assist in clarifying the requirements:

**File Information:** Participant Data, Created on July 10, 2017, File #1

**File Name:** CS1\_07102017\_PART\_001\_R01.CSV

**File Information:** Participant Data, Created on July 10, 2017, File #2

**File Name:** CS1\_07102017\_PART\_002\_R01.CSV

**File Information:** Benefits Data, Created on June 24, 2017, File #1  
**File Name:** CS1\_06242017\_BENE\_001\_R01.CSV

#### 4.2.8. Folder Structure

The data sharing device will be provided with an intact and project specific folder structure. There will be a top-level folder containing documentation or help files from the CAAL-Skills Pilot Project Team. Those folders will be created as Read-Only. The other top-level folder will be target location for the files generated by the Data Sharing Partners. This folder will be Read-Write). All files should be placed with the Partners folder without utilizing additional, supplemental, or sub-tier folders – that is, no folders should be added to the existing folder hierarchy.

**CAAL-Skills Folder (read only):** \CAAL-Skills

**Partner Source Data Folder (read-write):** \Data

## 4.3. Data Transfer

### 4.3.1. Introduction

The CAAL-Skills pilot project will facilitate data transfer from the source systems (Data Sharing Partner) to the target system (CAAL-Skills repository) utilizing project staff to physically distribute, collect, and secure the data transfer media. Future implementations of the data exchange process may involve electronic means. The data transfer media devices will be high capacity, encrypted Universal Serial Bus (USB) version 2.0 and 3.1 compatible, portable drives. The drives will include a physical encryption key which will encrypt the entire contents of the drive. The key will be unique for each data sharing partner – known only to the receiving staff and the CAAL-Skills pilot project architect and project manager. The devices will be secured on-site at the project office.

### 4.3.2. Device

The project will purchase and distribute to the Data Sharing Partners the Kingston DataTraveler 2000 USB drive also known as the DDT2000 for data transfer from the source system to the target system. The drive will have a capacity of 64GB. It is designed to be secure, with an alphanumeric keypad that locks the drive with a word or number combination, for easy-to-use PIN protection. Each Data Sharing Partner will have a unique PIN associated with their device. The DT2000 auto-lock feature is activated when the drive is removed from a device and it deletes the encryption key after 10 failed intrusion attempts. DT2000 features hardware-based, full disk AES 256-bit data encryption in XTS mode (see diagram below). Encryption is accomplished on the drive, not on the host computer system, and no trace of the PIN is left on the host system. DataTraveler 2000 is operating system independent and can be used on any device with a USB 2.0 or USB 3.1 Gen1 (USB 3.0) port. It's compatible with Windows®, Mac® OS, Linux, Chrome OS, and Android. The drive requires no additional software or drivers. It is FIPS 197 certified, to meet the most stringent IT security requirements for data encryption and transfer. The DT2000 enclosure is a durable design which protects the drive from everyday elements such as water and dust.



Complete specifications for the device are listed here:

- Interface: USB 3.1 Gen 1 (USB 3.0)
- Capacities: 16GB, 32GB & 64GB
- Speed: USB 3.1: 16GB: 120MB/s read, 20MB/s write 32GB: 135MB/s read, 40MB/s write 64GB: 135MB/s read, 40MB/s write



- USB 2.0: 16GB: 30MB/s read, 20MB/s write 32GB: 30MB/s read, 20MB/s write  
64GB: 30MB/s read, 20MB/s write
- Dimensions (drive with sleeve): 80mm x 20mm x 10.5mm
- Dimensions (drive without sleeve): 78mm x 18mm x 8mm
- Waterproof/Dustproof Certified: to IP57 rating
- Compatibility: USB 3.0 compliant and 2.0 compatible
- Operating Temperature: 0°C to 40°C
- Storage Temperature: -20°C to 60°C
- Minimum System Requirements: USB 3.0 compliant and 2.0 compatible
- Compatible with: Windows® 10, Windows 8.1, 8, 7 (SP1), Vista® (SP2), Mac OS X v. 10.9.x+, Linux v. 2.6.x+, Chrome OSTM, Android3

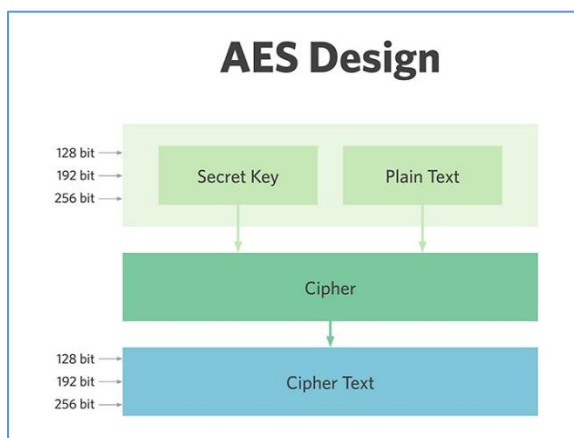
#### 4.3.3. Advanced Encryption Standard (AES)

The Advanced Encryption Standard (AES), also known by its original name Rijndael is a specification for the encryption of electronic data established by the U.S. National Institute of Standards and Technology (NIST) in 2001.

AES has been adopted by the U.S. government and is now used worldwide. It supersedes the Data Encryption Standard (DES), which was published in 1977. The algorithm described by AES is a symmetric-key algorithm, meaning the same key is used for both encrypting and decrypting the data.

In the United States, AES was announced by the NIST as U.S. FIPS PUB 197 (FIPS 197) on November 26, 2001. This announcement followed a five-year standardization process in which fifteen competing designs were presented and evaluated, before the Rijndael cipher was selected as the most suitable. AES became effective as a federal government standard on May 26, 2002, after approval by the Secretary of Commerce. AES is included in the ISO/IEC 18033-3 standard.

The Kingston Data Traveler DT2000 features hardware-based, full disk AES 256-bit data encryption in XTS mode. Encryption is done on the drive and no trace of the PIN is left on the system.



#### 4.3.4. Federal Information Processing Standards Certification

The Federal Information Processing Standards (FIPS) are U.S. government standards for information technology and computer security. The FIPS program is run by the National Institute of Standards (NIST). NIST FIPS 140 is the cryptography standard program required by the US federal government for protection of sensitive data. All products using cryptography in use by civilian and military US federal agencies must comply with FIPS 140. The current FIPS version is 140-2.

Obtaining FIPS 140 validation requires a rigorous testing process by an accredited testing laboratory. The results are then reviewed by the National Institute of Standards and Technology (NIST) which issues a FIPS 140 validation certificate. The certificates and policies for the DT2000 generation 1 and 2 devices are listed below:

Product Name	DT2000 Gen2	DT2000 Gen1
Firmware Version	3.05	3.05
FIPS certificate	#2307	#2371
Security Policy	#2307	#2371
FIPS AES certificate	#2838	#2838
FIPS SHS certificate	#2379	#2379
FIPS RSA certificate	#1480	#1480
FIPS DRBG certificate	#494	#494
FIPS RNG certificate		
FIPS HMAC certificate	#1779	#1779

#### 4.3.5. Hardware Encryption

Hardware-based encryption as built into the device indicated above is notably transparent to the user. The symmetric encryption key is maintained independently from the CPU, thus removing computer memory as a potential attack vector. The drive except for boot-up authentication operates just like any drive with no degradation in performance. There is no complication or performance overhead, unlike disk encryption software, since all the encryption is invisible to the operating system and the host computers processor.

Hardware based encryption offers these benefits:

- Uses a dedicated processor physically located on the encrypted drive
- Processor contains a random number generator to generate an encryption key, which the user's password will unlock
- Increased performance by off-loading encryption from the host system

- Safeguard keys and critical security parameters within crypto-hardware
- Authentication takes place on the hardware
- Encryption is tied to a specific device, so encryption is “always on”
- Does not require any type of driver installation or software installation on host PC
- Protects against the most common attacks, such as cold boot attacks, malicious code, brute force attack

#### 4.3.6. Device Format

The Kingston Data Traveler 2000 USB encrypted drive will be formatted in a Windows NTFS 5.0 or later file system. This system is compatible with Windows 7, 8, 10, Windows Server editions, as well as Ubuntu and Red Hat Linux 6.x. The file system will not be encrypted (encryption will be applied at the hardware level) and will not be compressed. The file system should not be altered nor the device re-formatted by the Data Sharing Partners. NTFS 5 was chosen over FAT32 and FAT16 due to the improved file handling capabilities and capacities as noted below:

Criteria	NTFS5	FAT32	FAT16
<b>Operating System Capability</b>	Windows 2000 Windows XP Windows 2003/5/8/12/16 Server Windows Vista Windows 7 Windows 8 Windows 10	DOS v7 and higher Windows 98 Windows ME Windows 2000 Windows XP Windows 2003 Server Windows Vista Windows 7	DOS All versions of Microsoft Windows
<b>Max Volume Size</b>	$2^{64}$ clusters – 1 cluster	32GB for all OS. 2TB for some OS	2GB for all OS. 4GB for some OS
<b>Max Files on Volume</b>	4,294,967,295 $2^{32} - 1$	4194304	65536
<b>Max File Size</b>	$2^{64}$ bytes (16 ExaBytes) minus 1KB	4GB minus 2 Bytes	2GB (Limit Only by Volume Size)
<b>Max Clusters Number</b>	$2^{64}$ clusters – 1 cluster	4177918	65520
<b>Max File Name Length</b>	Up to 255	Up to 255	Standard - 8.3 Extended - up to 255

<b>Unicode File Names</b>	Unicode Character Set	System Character Set	System Character Set
<b>System Records Mirror</b>	MFT Mirror File	Second Copy of FAT	Second Copy of FAT
<b>Boot Sector Location</b>	First and Last Sectors	First Sector and Copy in Sector #6	First Sector
<b>File Attributes</b>	Standard and Custom	Standard Set	Standard Set
<b>Alternate Streams</b>	Yes	No	No
<b>Compression</b>	Yes	No	No
<b>Encryption</b>	Yes	No	No
<b>Object Permissions</b>	Yes	No	No
<b>Disk Quotas</b>	Yes	No	No
<b>Sparse Files</b>	Yes	No	No
<b>Reparse Points</b>	Yes	No	No
<b>Volume Mount Points</b>	Yes	No	No
<b>Built-In Security</b>	Yes	No	No
<b>Recoverability</b>	Yes	No	No
<b>Performance</b>	Low on small volumes High on Large	High on small volumes Low on large	Highest on small volumes Low on large
<b>Disk Space Economy</b>	Max	Average	Minimal on large volumes
<b>Fault Tolerance</b>	Max	Minimal	Average

#### 4.3.7. Device Identification

Each Kingston Data Traveler DT2000 will be serialized with the serial number appearing in a 24 character, 2D bar-code readable element on the case. The serial numbers will be unique and correspond to each Data Sharing Partner. Also, the drive will include a folder titled “\CAAL-Skills” containing the name of the Data Sharing Partner. The CAAL-Skills project team will maintain an inventory of the devices, serial numbers, assignments (to Data Sharing Partners) and a complete and current log of the chain of custody. Finally, the devices will be stored in a locked, fireproof safe in the CWDB facilities.

#### 4.3.8. Device Loading

The actual steps for placing the CAAL-Skills formatted files on the device from the source system are a combination of steps and processes from the Data Sharing Partner’s Information Technology (DSP-IT) transfer standards and the CAAL-Skills pilot project. In the steps below, those where the owner/responsible party are identified as “*DSP-IT proprietary*” are recommendations only and should conform to the DSP-IT data transfer standards, policies, and procedures.

1. Remove the protective cover from the Kingston Data Traveler 2000 USB encrypted drive.
2. Enter the assigned PIN (unique to each Data Sharing Partner and noted on the chain of custody form) via the keypad on the Kingston Data Traveler 2000 encrypted drive to unlock the drive and decrypt the contents
3. Insert the Kingston Data Traveler 2000 drive into a desktop or server with access to the CAAL-Skills formatted files.
4. Copy all requested files into the \Data folder on the Kingston Data Traveler 2000 drive. *DSP-IT proprietary.*
5. Using the operating system appropriate procedures eject the device from the USB port to ensure the files are written and properly closed. *DSP-IT proprietary.*
6. Upon removal, press the “Lock” button on the Kingston Data Traveler 2000 drive to encrypt the contents.
7. Insert the Kingston Data Traveler 2000 encrypted drive into the sleeve and follow the Chain of Custody process.

#### 4.3.9. Device Chain of Custody

Chain of custody defines the processes, procedures, and forms to manage, track, and document the transfer of the CAAL-Skills pilot project requested information from the Data Sharing Partners. The chain includes: the initial transfer of the encrypted data device from CAAL-Skills to the partner, the return of the device with the requested data, and the receipt of the device and internal secured storage of the device within the CAAL-Skills project team. The standard for this project will be based on the United States Department of Commerce’s National Institute of Standards and Technology (NIST) “Computer Security Handling Guide” Special Publication 801-

61 Revision 2. Specifically, this document will reference the section “Coordination and Information Sharing” and the NIST Chain of Custody Tracking Form. As shown below:

<b>Data Transfer Device #:</b> _____																																																		
<b>CAAL-Skills Pilot Project</b> <b>DATA DEVICE CHAIN OF CUSTODY TRACKING FORM</b>																																																		
Data Sharing Partner (agency): _____ Location of Device (address): _____ Primary Contact Name: _____ Primary Contact Phone: _____ Primary Email Address: _____ Alternate Contact Name: _____ Alternate Contact Phone: _____ Alternate Email Address: _____																																																		
<b>Description of Device</b>																																																		
<table border="1" style="width: 100%; border-collapse: collapse;"><thead><tr><th style="width: 10%;">Item #</th><th style="width: 10%;">Quantity</th><th style="width: 80%;">Description of Item (Model, Serial #, External Condition)</th></tr></thead><tbody><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></tbody></table>	Item #	Quantity	Description of Item (Model, Serial #, External Condition)																																															
Item #	Quantity	Description of Item (Model, Serial #, External Condition)																																																
<b>Chain of Custody</b>																																																		
<table border="1" style="width: 100%; border-collapse: collapse;"><thead><tr><th style="width: 10%;">Item #</th><th style="width: 15%;">Date/Time</th><th style="width: 25%;">Released by Data Sharing Partner (Print and Sign)</th><th style="width: 25%;">Received by CAAL-Skills Staff (Print and Sign#)</th><th style="width: 25%;">Comments/Location</th></tr></thead><tbody><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td><td> </td><td> </td></tr></tbody></table>	Item #	Date/Time	Released by Data Sharing Partner (Print and Sign)	Received by CAAL-Skills Staff (Print and Sign#)	Comments/Location																																													
Item #	Date/Time	Released by Data Sharing Partner (Print and Sign)	Received by CAAL-Skills Staff (Print and Sign#)	Comments/Location																																														
<div style="display: flex; justify-content: space-between;"><span>APD_Form_#PE003_v.1 (12/2013)</span><span>Page 1 of 2 pages</span></div> <div style="margin-top: 10px; font-size: small;"><p>Technical Working Group on Information Technology Evidence Preservation: <i>Best Practices for Data Exchange Partners</i>. U.S. Department of Commerce, National Institute of Standards and Technology. 2013.</p></div>																																																		

**DATA DEVICE CHAIN-OF-CUSTODY TRACKING FORM**  
(Continued)

Chain of Custody				
Item #	Date/Time	Released by Data Sharing Partner (Print and Sign)	Received by CAAL-Skills (Print and Sign#)	Comments/Location



**Final Disposal Authority**

**Authorization for Disposal**

Item(s) #: \_\_\_\_\_ on this document pertaining to (suspect): \_\_\_\_\_  
is/are no longer needed for information transfer and is/are authorized for disposal by (check appropriate disposal method)

☐ Return to Owner      ☐ Auction/Destroy/Divert

Name of Authorizing Owner: \_\_\_\_\_ Signature: \_\_\_\_\_ Date: \_\_\_\_\_

**Witness to Destruction of Evidence**

Item(s) #: \_\_\_\_\_ on this document were destroyed by Evidence Custodian \_\_\_\_\_  
in my presence on (date) \_\_\_\_\_  
Witness to destruction: \_\_\_\_\_ Signature: \_\_\_\_\_ Date: \_\_\_\_\_

This Evidence Chain-of-Custody form is to be retained as a permanent record by the CAAL-Skills project and California Workforce Development Board staff.

Essentially, whenever the data device changes ownership, that change should be reflected in the above form.

## 5. Other Project Agreements

### 5.1. Data Sharing Agreements

Each Data Sharing Partner will be party to a Data Sharing Agreement which outlines the terms, scope, costs, and timing of source data extracts to be provided to the CAAL-Skills pilot project. While the details of those individual agreements are being addressed, the project will work directly with the Partner's assigned information technology staff to design and test the source data extraction processes. Once the agreements have been signed by all parties, the source data will be officially requested.



## 6. Partner Response

### 6.1. High Level Tasks and Key Dates

The following table is to be completed by the Data Sharing Partner to assist the CAAL-Skills project team in determining overall project timelines for receiving and transforming the data.

ID	Task	Duration	Start Date	End Date	Notes
1	Identify Data Extraction Resources				
2	Quantify Cost / Time for Data Extraction				
3	Engage Resources for Data Extraction				
4	Generate and Transfer Test Data Files				
5	Implement Modifications (if necessary)				
6	Generate and Transfer Final Data Files				

### 6.2. High Level Tasks and Costs

The following table is to be completed by the Data Sharing Partner to assist the CAAL-Skills project team in determining costs for each partner to design systems and generate the required data files for use by the CAAL-Skills team.

ID	Task	Costs	Notes
1	Identify Data Extraction Resources		
2	Quantify Cost / Time for Data Extraction		
3	Engage Resources for Data Extraction		
4	Generate and Transfer Test Data Files		
5	Implement Modifications (if necessary)		
6	Generate and Transfer Final Data Files		

### 6.3. Response Timeframe

The CAAL-Skills Pilot Project team is requesting a response from the Partners within two weeks of reviewing this document with the CAAL-Skills team. The response should include estimated dates and costs for key tasks as well as primary and alternate contact information (see next section). The response should be emailed to the CAAL-Skills Pilot Project Manager, Patrick Getz, [patrick.getz@cwdb.ca.gov](mailto:patrick.getz@cwdb.ca.gov)

## 7. Contact Information

### 7.1. CAAL-Skills Project – Chief Technical Architect

**Paul Banas**  
**California Workforce Development Board**  
**800 Capitol Mall, Ste, 1022**  
**Sacramento, CA 95814**  
**(916) 657-1483 desk**  
**(916) 657-1381 fax**  
**(616) 206-0727 mobile**  
[paul.banas@cwddb.ca.gov](mailto:paul.banas@cwddb.ca.gov)

### 7.2. Primary Technical Contact

Name: \_\_\_\_\_  
Title: \_\_\_\_\_  
Agency/Department: \_\_\_\_\_  
Address: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
Desk phone: \_\_\_\_\_  
Cell phone: \_\_\_\_\_  
Email address: \_\_\_\_\_

### 7.3. Alternate Technical Contact

Name: \_\_\_\_\_  
Title: \_\_\_\_\_  
Agency/Department: \_\_\_\_\_  
Address: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
Desk phone: \_\_\_\_\_  
Cell phone: \_\_\_\_\_  
Email address: \_\_\_\_\_

## 8. Data Element Dictionary

### 8.1. Introduction

Each data element will be presented on a separate page with the following attributes:

1. DED# - Project unique Data Element Dictionary number
2. Data Element Name – Descriptive name of data element
3. Format – the element’s data type, length, precision, and scale
4. Definition – the full definition of the element
5. CAAL-Skills Field Name – the field to be inserted in the header row of the data exchange file
6. CAAL-Skills Values – the expected values for the element. In some cases, there will be a translation or cross reference table listed, others will note the “actual source value” in which case the source system’s fully qualified value should be inserted. Note: fully qualified = the full descriptive name rather than the source system code. *For example: a valid value for Ethnicity/Race would be “Hispanic / Latino”, not the internal code or representation “001”.*
7. CAAL-Skills File Type – the name of the File grouping identified in section 4.2.7.
8. Collection Dates – The date range the element should be collected and provided to CAAL-Skills

DED#	Data Element Name	Format
005	Social Security Number	INT
<b>Definition</b>		
The social security number assigned to the participant.		
<b>CAAL Skills Field Name</b>		
SSN		
<b>CAAL Skills Values</b>		
Actual source value		
97 = Other		
98 = Not Applicable		
99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

Diagram illustrating the structure of the Data Element Dictionary table with numbered boxes 1 through 8 pointing to specific attributes:

- 1 points to DED#
- 2 points to Data Element Name
- 3 points to Format
- 4 points to Definition
- 5 points to CAAL Skills Field Name
- 6 points to CAAL Skills Values
- 7 points to File
- 8 points to Collection Dates

## Partner Identification

DED#	Data Element Name	Format
000	Partner ID	INT
<b>Definition</b>		
The internally assigned two-digit code for each Data Sharing Partner / Administrative Agency.		
<b>CAAL Skills Field Name</b>		
AAID		
<b>CAAL Skills Values</b>		
1 - California Department of Education 2 - California Department of Rehabilitation 3 - California Department of Industrial Relations 4 - California Community Colleges Chancellor's Office 5 - California Department of Social Services 6 - Employment Training Panel 7 - California Employment Development Department – Central Office Workforce Services Division 8 - California Employment Development Department – Labor Market Information Division 9 - California Employment Development Department – Tax Branch 10 - CWDA Napa County 11 - CWDA Monterey County 12 - CWDA Stanislaus County 97 = Other 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Administrative Agency / Partner		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

## Enrolling Organization

DED#	Data Element Name	Format
002	Enrollment Organization Name	VCHAR(40)
<b>Definition</b>		
The name of the organization where the participant was determined eligible for the program and was enrolled to receive their first service financially assisted by the program.		
<b>CAAL Skills Field Name</b>		
EONM		
<b>CAAL Skills Values</b>		
For ETP, this value will be "ETP" For CCCCCO, this value will be the name of the community college For EDD Workforce Services Division, this value will be the name of the local workforce board For CWDA-Counties, this value will be the county name For all others, the value will be the actual source system name (enrollment organization name) 99 = Unknown		
<b>File</b>		
Enrolling Organization		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
003	Enrollment Organization ID	VCHAR(40)
<b>Definition</b>		
The identifier assigned to the organization where the participant was determined eligible for the program and was enrolled to receive their first service financially assisted by the program.		
<b>CAAL Skills Field Name</b>		
EOID		
<b>CAAL Skills Values</b>		
For CWDA-Counties, this value will be the county code (e.g. Napa = 28, Monterey = 27, Stanislaus = 50) For all others, actual source system ID 99 = Unknown		
<b>File</b>		
Enrolling Organization		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

## Funding Source

DED#	Data Element Name	Format
004	Fund Source	INT
<b>Definition</b>		
The funding source used to pay for the services received by the participant.		
<b>CAAL Skills Field Name</b>		
FSRC		
<b>CAAL Skills Values</b>		
1 = WIOA Title I (Adult) 2 = WIOA Title I (Dislocated Worker) 3 = WIOA Title I (Youth) 4 = WIOA Title II 5 = WIOA Title III 6 = WIOA Title IV 7 = CalWORKs 8 = CalFresh 9 = TAA 97 = Other 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Services		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

## Participant Identification

DED#	Data Element Name	Format
005	Social Security Number	INT
<b>Definition</b>		
The social security number assigned to the participant.		
<b>CAAL Skills Field Name</b>		
SSN		
<b>CAAL Skills Values</b>		
Actual source value 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		



DED#	Data Element Name	Format
006	Participant First Name	VCHAR(40)
<b>Definition</b>		
The first name of the participant.		
<b>CAAL Skills Field Name</b>		
PFNM		
<b>CAAL Skills Values</b>		
Actual source value 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
007	Participant Middle Name	VCHAR(40)
<b>Definition</b>		
The middle name or initial of the participant.		
<b>CAAL Skills Field Name</b>		
PMNM		
<b>CAAL Skills Values</b>		
Actual source value 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
008	Participant Last Name	VCHAR(40)
<b>Definition</b>		
The last name of the participant.		
<b>CAAL Skills Field Name</b>		
PLNM		
<b>CAAL Skills Values</b>		
Actual source value 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
010	Participant City	VCHAR(50)
<b>Definition</b>		
The city associated to the physical, mailing, or billing address of the participant.		
<b>CAAL Skills Field Name</b>		
PCITY		
<b>CAAL Skills Values</b>		
Actual source value 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
011	Participant Postal Code	VCHAR(9)
<b>Definition</b>		
The postal code associated to the physical, mailing, or billing address of the participant. For US Zip Codes, include “+4” numeric digits if available. Exclude the “+” notation.		
<b>CAAL Skills Field Name</b>		
PPOST		
<b>CAAL Skills Values</b>		
Actual source value 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
012	Participant State Code	VCHAR(2)
<b>Definition</b>		
The two letter state code associated to the physical, mailing, or billing address of the participant for US addresses. See Skills Values for all other addresses.		
<b>CAAL Skills Field Name</b>		
PST		
<b>CAAL Skills Values</b>		
Actual source value CN = Canada MX = Mexico 97 = All other countries 98 = Not Applicable 99 = Unknown AE = (ZIPs 09xxx) Armed Forces Europe including Canada, Middle East, and Africa AP = (ZIPs 962xx - 966xx) for Armed Forces Pacific AA (ZIPs 340xx) for Armed Forces (Central and South) Americas		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
013	Participant Date of Birth	DATE
<b>Definition</b>		
The month, day, and year the participant was born.		
<b>CAAL Skills Field Name</b>		
DOB		
<b>CAAL Skills Values</b>		
Actual source value 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
014	Participant Sex / Gender	VCHAR(2)
<b>Definition</b>		
The sex the participant identifies as.		
<b>CAAL Skills Field Name</b>		
SEX		
<b>CAAL Skills Values</b>		
M = Male F = Female 97 = Other 99 = Unknown or not provided		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		



DED#	Data Element Name	Format
015	Participant Unique Individual Identifier	VCHAR(20)
<b>Definition</b>		
The unique identifier assigned to the participant, by an organization, to distinguish the participant from other participants. This value may be the internal system identifier, a surrogate key assigned at time of data extract, or a concatenated natural key. The value must be consistent for all extracts to ensure proper identification.		
<b>CAAL Skills Field Name</b>		
PUID		
<b>CAAL Skills Values</b>		
For CWDA-Counties, this value will be the Client Identification Number or CIN For all others, actual source value 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

## Participant Demographics

DED#	Data Element Name	Format
017	Participant Ethnicity	VCHAR(100)
<b>Definition</b>		
The ethnicity the participant identifies as.		
<b>CAAL Skills Field Name</b>		
ETHN		
<b>CAAL Skills Values</b>		
Actual source values, i.e. not internal system codes or numeric representations. Please use the fully qualified value. For multiple values, use comma separation.		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
017-1	Participant Race	VCHAR(100)
<b>Definition</b>		
The race the participant identifies as.		
<b>CAAL Skills Field Name</b>		
RACE		
<b>CAAL Skills Values</b>		
Actual source values, i.e. not internal system codes or numeric representations. Please use the fully qualified value. For multiple values, use comma separation.		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
018	Participant Veteran Status	INT
<b>Definition</b>		
If the participant identifies as a veteran of the armed forces.		
<b>CAAL Skills Field Name</b>		
VET		
<b>CAAL Skills Values</b>		
1 = Yes 0 = No 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
020	Homeless Individual	INT
<b>Definition</b>		
If the participant identifies as a homeless individual at program entry.		
<b>CAAL Skills Field Name</b>		
HMLS		
<b>CAAL Skills Values</b>		
1 = Yes 0 = No 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
023	Participant Foster Care Youth	INT
<b>Definition</b>		
If the participant identifies as a current / former foster care youth.		
<b>CAAL Skills Field Name</b>		
FSTRC		
<b>CAAL Skills Values</b>		
1 = Yes 0 = No 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
027	Temporary / Permanent Disability	INT
<b>Definition</b>		
If the participant identifies as, or was determined by authorized personnel as, having a disability.		
<b>CAAL Skills Field Name</b>		
DISA		
<b>CAAL Skills Values</b>		
1 = Yes, Disabled (For CWDA-Counties, Exemption Code 600) 2 = Yes, Temporarily Disabled (For CWDA-Counties, Exemption Code 302) 0 = No 98 = Not Applicable 99 = Unknown (For CWDA-Counties, No Exemption Code or Codes other than 600 and 302)		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
028	Migrant / Seasonal Farmworker	INT
<b>Definition</b>		
If the participant identifies as a migrant / seasonal farmworker or a child of a migrant / seasonal farmworker.		
<b>CAAL Skills Field Name</b>		
FARM		
<b>CAAL Skills Values</b>		
1 = Yes 0 = No 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		



DED#	Data Element Name	Format
029	Participant Single Parent	INT
<b>Definition</b>		
If the participant identifies as a single parent. For CWDA-Counties the definition is expanded: "Single parent may include an individual who is the sole parent of an individual who is the sole provider in a family, as their spouse/partner is disabled."		
<b>CAAL Skills Field Name</b>		
SPAR		
<b>CAAL Skills Values</b>		
1 = Yes 0 = No 98 = Not Applicable 99 = Unknown		
<b>File</b>		
Participants		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

## Workforce Training and Related Education

DED#	Data Element Name	Format
035	Type of Career Service	INT
<b>Definition</b>		
The type of career service assigned to and/or received by the participant.		
<b>CAAL Skills Field Name</b>		
TCS		
<b>CAAL Skills Values</b>		
1 - Basic Career Service (Staff Assisted) 2 - Individualized Career Service 3 - Information Only Career Service 4 - Follow up Career Service 5 - Basic Career Service (Self-Service) 6 - Vocational Rehabilitation Service 97 = Other 98 = Not Applicable 99 = Unknown  *Actual source values collected for CCCCCO, CWDA County Partners, and DOR.  ** CWDA County Partner Examples: 1) CalWORKS Activity – Appraisal 2) CalWORKS Activity – Assessment 3) CalWORKS Activity – Reappraisal 4) CalWORKS Activity – Job search & job readiness assistance 5) CalWORKS Activity - Unsubsidized employment 6) CalWORKS Activity - Self-employment 7) CalWORKS Activity - Subsidized private sector employment 8) CalWORKS Activity - Subsidized public sector employment 9) CalWORKS Activity - On-the-job training (OJT) 10) CalWORKS Activity - Grant-based on-the-job training (OJT) 11) CalWORKS Activity - Work-study 12) CalWORKS Activity - Supported work or transitional employment 13) CalWORKS Activity - Work experience 14) CalWORKS Activity - Community service 15) CalWORKS Activity - Job skills training directly related to employment 16) CalWORKS Activity - Vocational education and training 17) CalWORKS Activity - Education directly related to employment 18) CalWORKS Activity - Adult basic education 19) CalWORKS Activity - Satisfactory progress in a secondary school 20) CalWORKS Activity - Other activities 21) CalWORKS Activity - Providing childcare to community services participants 22) CalWORKS Activity - Mental health services		

- 23) CalWORKS Activity - Domestic abuse services
- 24) CalWORKS Activity - Post-employment/job retention services
- 25) CalWORKS Activity - Post CalWORKs 48-month service
- 26) CalWORKS Service – Transportation
- 27) CalWORKS Service – Ancillary
- 28) CalWORKS Service – Childcare
  
- 1) CalFresh Activity - Job search
- 2) CalFresh Activity - Job search training
- 3) CalFresh Activity - Workfare
- 4) CalFresh Activity - Work experience
- 5) CalFresh Activity - Education
- 6) CalFresh Activity - Self-employment training
- 7) CalFresh Activity - WIOA Activities
- 8) CalFresh Service - Transportation/ancillary
- 9) CalFresh Service - Dependent care

<b>File</b>
Services
<b>Collection Dates</b>
7/1/13 – 6/30/17

DED#	Data Element Name	Format
038	Employment and Training Service Related to Supplemental Nutritional Assistance Program (SNAP E&T) or Welfare to Work	INT
<b>Definition</b>		
If the participant received employment and training services from the Supplemental Nutritional Assistance Program (SNAP) or the Welfare to Work Program.		
<b>CAAL Skills Field Name</b>		
ETSR		
<b>CAAL Skills Values</b>		
1 = Yes, SNAP E&T 2 = Yes, Welfare to Work 3 = No Leave Blank = Unknown  Note: if a participant is eligible to receive services under both programs then the Data Element Format will be VCHAR(5) with a comma separated list of values.		
<b>File</b>		
Services		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
036	Type of Training Service	INT
<b>Definition</b>		
The type of training service assigned to and/or received by the participant.		
<b>CAAL Skills Field Name</b>		
TTS		
<b>CAAL Skills Values *</b>		
0 = No Training Service 1 = On the Job Training 2 = Skill Upgrading 3 = Entrepreneurial Training 4 = ABE or ESL in conjunction with Training (non-TAA funded) 5 = Customized Training 6 = Other Occupational Skills Training 7 = Remedial Training (ABE/ESL – TAA only) 8 = Prerequisite Training 9 = Apprenticeship Training 10 = Other Basic Skills Training (WIA Youth) 11 = Occupational Skills Training (Youth) 12 = Other Non-Occupational-Skills Training 13 - Unspecified Training 97 = Other 98 = Not Applicable 99 = Unknown  *Actual source value for DOR and CWDA County Partners  ** CWDA County Partner Examples: 1) CalWORKS Activity – Appraisal 2) CalWORKS Activity – Assessment 3) CalWORKS Activity – Reappraisal 4) CalWORKS Activity – Job search & job readiness assistance 5) CalWORKS Activity - Unsubsidized employment 6) CalWORKS Activity - Self-employment 7) CalWORKS Activity - Subsidized private sector employment 8) CalWORKS Activity - Subsidized public sector employment 9) CalWORKS Activity - On-the-job training (OJT) 10) CalWORKS Activity - Grant-based on-the-job training (OJT) 11) CalWORKS Activity - Work-study 12) CalWORKS Activity - Supported work or transitional employment 13) CalWORKS Activity - Work experience 14) CalWORKS Activity - Community service 15) CalWORKS Activity - Job skills training directly related to employment 16) CalWORKS Activity - Vocational education and training		

- 17) CalWORKS Activity - Education directly related to employment
  - 18) CalWORKS Activity - Adult basic education
  - 19) CalWORKS Activity - Satisfactory progress in a secondary school
  - 20) CalWORKS Activity - Other activities
  - 21) CalWORKS Activity - Providing childcare to community services participants
  - 22) CalWORKS Activity - Mental health services
  - 23) CalWORKS Activity - Domestic abuse services
  - 24) CalWORKS Activity - Post-employment/job retention services
  - 25) CalWORKS Activity - Post CalWORKs 48-month service
  - 26) CalWORKS Service – Transportation
  - 27) CalWORKS Service – Ancillary
  - 28) CalWORKS Service – Childcare
- 
- 1) CalFresh Activity - Job search
  - 2) CalFresh Activity - Job search training
  - 3) CalFresh Activity - Workfare
  - 4) CalFresh Activity - Work experience
  - 5) CalFresh Activity - Education
  - 6) CalFresh Activity - Self-employment training
  - 7) CalFresh Activity - WIOA Activities
  - 8) CalFresh Service - Transportation/ancillary
  - 9) CalFresh Service - Dependent care

<b>File</b>
Services
<b>Collection Dates</b>
7/1/13 – 6/30/17

## Supportive Services

DED#	Data Element Name	Format
064	Type of Supportive Service or Activity	VCHAR(80) Comma Separated List
<b>Definition</b>		
The type of supportive service or activity assigned to and/or received by the participant.		
<b>CAAL Skills Field Name</b>		
TOSS		
<b>CAAL Skills Values *</b>		
Actual source values including "CalWORKs" and "CalFRESH" prefix **		
97 = Other		
99 = Unknown		
*Actual source values collected for CWDA County Partners and DOR.		
** CWDA County Partner Examples:		
1) CalWORKS Activity – Appraisal		
2) CalWORKS Activity – Assessment		
3) CalWORKS Activity – Reappraisal		
4) CalWORKS Activity – Job search & job readiness assistance		
5) CalWORKS Activity - Unsubsidized employment		
6) CalWORKS Activity - Self-employment		
7) CalWORKS Activity - Subsidized private sector employment		
8) CalWORKS Activity - Subsidized public sector employment		
9) CalWORKS Activity - On-the-job training (OJT)		
10) CalWORKS Activity - Grant-based on-the-job training (OJT)		
11) CalWORKS Activity - Work-study		
12) CalWORKS Activity - Supported work or transitional employment		
13) CalWORKS Activity - Work experience		
14) CalWORKS Activity - Community service		
15) CalWORKS Activity - Job skills training directly related to employment		
16) CalWORKS Activity - Vocational education and training		
17) CalWORKS Activity - Education directly related to employment		
18) CalWORKS Activity - Adult basic education		
19) CalWORKS Activity - Satisfactory progress in a secondary school		
20) CalWORKS Activity - Other activities		
21) CalWORKS Activity - Providing childcare to community services participants		
22) CalWORKS Activity - Mental health services		
23) CalWORKS Activity - Domestic abuse services		
24) CalWORKS Activity - Post-employment/job retention services		
25) CalWORKS Activity - Post CalWORKs 48-month service		
26) CalWORKS Service – Transportation		
27) CalWORKS Service – Ancillary		
28) CalWORKS Service – Childcare		

- 1) CalFresh Activity - Job search
- 2) CalFresh Activity - Job search training
- 3) CalFresh Activity - Workfare
- 4) CalFresh Activity - Work experience
- 5) CalFresh Activity - Education
- 6) CalFresh Activity - Self-employment training
- 7) CalFresh Activity - WIOA Activities
- 8) CalFresh Service - Transportation/ancillary
- 9) CalFresh Service - Dependent care

<b>File</b>
Supportive Services
<b>Collection Dates</b>
7/1/13 – 6/30/17



DED#	Data Element Name	Format
065	First Date of Service or Activity	DATE
<b>Definition</b>		
The month, day, and year the participant began receiving a specified service or activity.		
<b>CAAL Skills Field Name</b>		
FDSS		
<b>CAAL Skills Values</b>		
Actual source values		
<b>File</b>		
Supportive Services		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		

DED#	Data Element Name	Format
066	Last Date of Service or Activity	DATE
<b>Definition</b>		
The month, day, and year the participant stopped receiving a specified service or activity.		
<b>CAAL Skills Field Name</b>		
LDSS		
<b>CAAL Skills Values</b>		
Actual source values		
<b>File</b>		
Supportive Services		
<b>Collection Dates</b>		
7/1/13 – 6/30/17		