# Human Services Data Specification

This specification defines a minimum set of data elements for a human services data specification. The idea behind the data specification is to create data sets that can be used by multiple domains of expertise. The goal is to derive a specification robust enough to support Information and Retrieval (IR) applications and flexible enough to support localized directories. The focus of this specification is content but a large portion of the specification is dedicated to format because reducing complex data structures into a commonly used format such as comma delimited files (CSV) requires decomposing them into separate files.

# Term Definitions

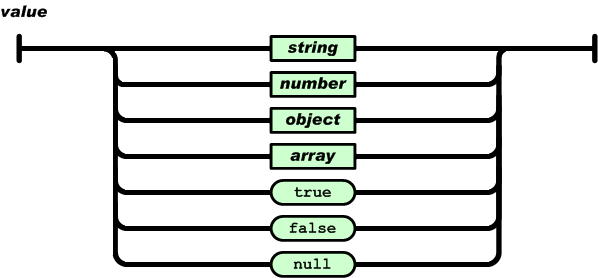
* Field required - The field column must be included in your feed, and a value must be provided for each record. Some required fields permit an empty string as a value. To enter an empty string, just omit any text between the commas for that field. Note that 0 is interpreted as “a string of value 0”, and is not an empty string. Please see the field definition for details.
* Field optional - The field column may be omitted from your feed. If you choose to include an optional column, each record in your feed must have a value for that column. You may include an empty string as a value for records that do not have values for the column. Some optional fields permit an empty string as a value. To enter an empty string, just omit any text between the commas for that field. Note that 0 is interpreted as “a string of value 0”, and is not an empty string.

# Conventions

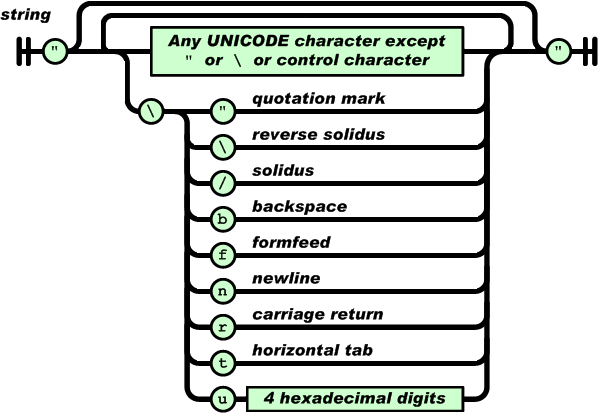
This specification uses snake case where compound words are joined by an underscore. Field name are all lower case and in snake case, e.g. ‘organization\_id’.

# Data Types

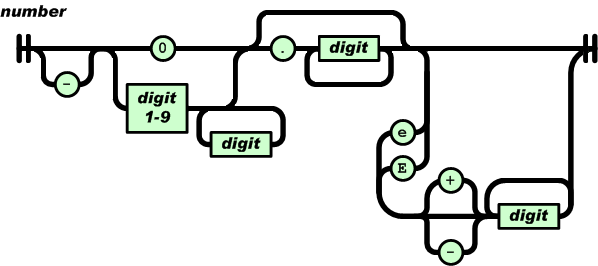
Data types are borrowed from JSON (JavaScript Object Notation). The types represented in JSON are strings, numbers, objects, arrays, booleans, and null.

 “Data types”

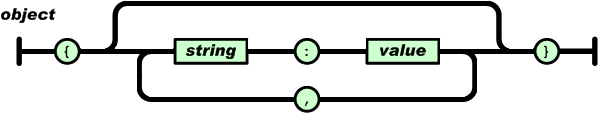
A string is a sequence of zero or more characters wrapped in quotes with backslash escapement.

 “String”

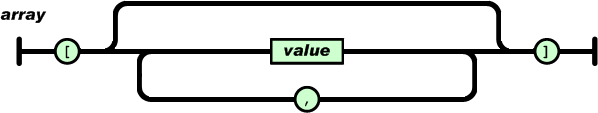
A number can be represented as integer, real, or floating point. JSON does not support octal or hex because it is minimal. It does not have values for NaN or Infinity. Numbers are not quoted.

 “Number”

A JSON object is an unordered collection of key/value pairs wrapped in curly braces. The keys are strings and the values are any of the JSON types. A colon separates the keys from the values, and comma separates the pairs.

 “Object”

The JSON array is an ordered collection of values separated by commas wrapped in square brackets.

 “Array”

JSON has two boolean values, true and false.

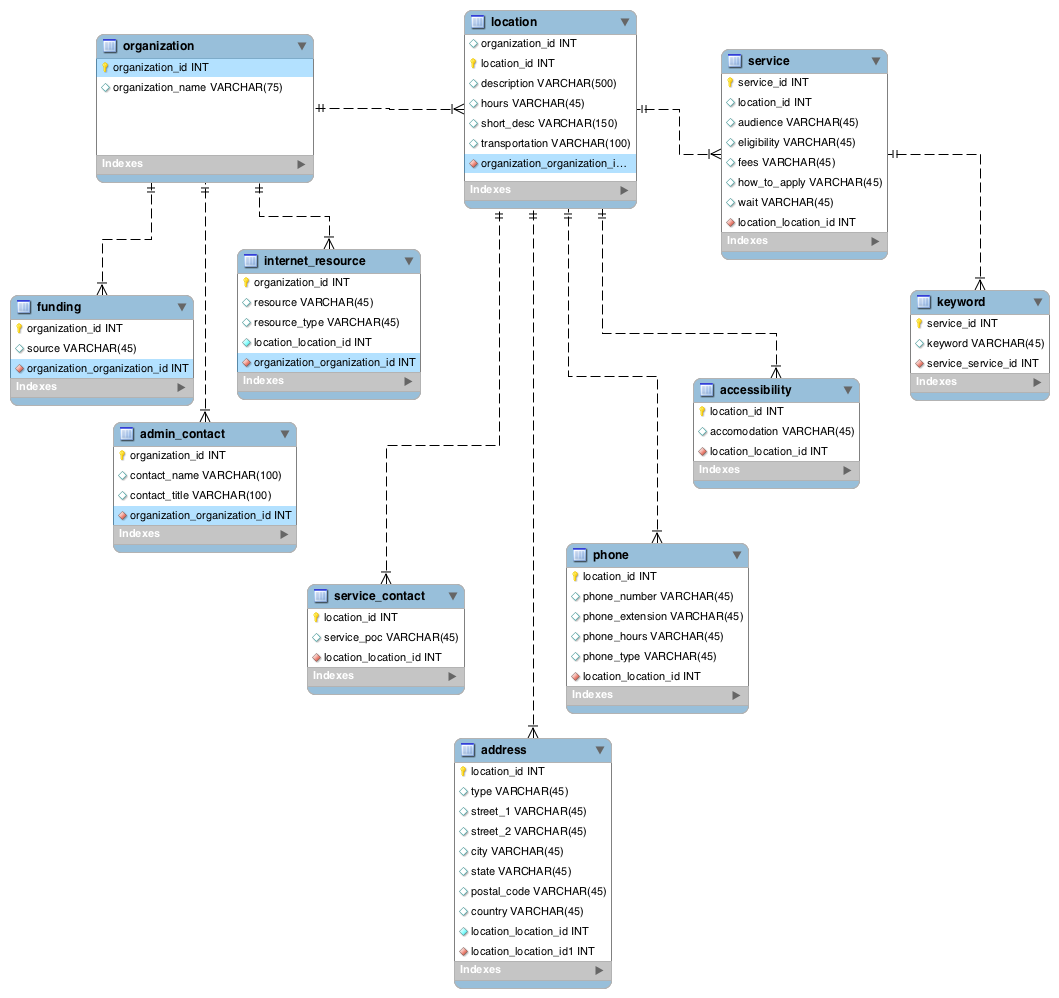
# Data Model

The basic entities are an organization or agency, a program or service, and a site or location. The structure is hierarchical in that an organization has programs that have locations. Note that an organization can have one or more locations and locations can have one or more programs.

The data model can be expressed as JSON. Note that this is an application specific way to express the data, it can be expressed to fit other schemas or application models.

[  
 {  
 "name": "Mid-Peninsula Speech and Language Clinic",  
 "locations": [  
 {  
 "name": "Mid-Peninsula Speech and Language Clinic",  
 "contacts": [  
 {  
 "contact\_name": "Donna Dagenais, MA",  
 "contact\_title": "Director"  
 }  
 ],  
 "description": "Provides evaluation, therapy and follow-up to persons of any age who have speech, language or learning disorders. Speech and language disorders may include articulation problems, cleft palate, voice disorders, stuttering, speech retraining following stroke or neurological trauma or neurologic disease. Individuals are first seen for evaluation, after which recommendations are made regarding appropriate treatment or referral for other professional services. Suggestions can be made to help design individual education plans for children needing special services in school, individual and group treatment, tutorial assistance and family counseling.",  
 "short\_desc": "Provides evaluation, therapy and follow-up to persons of any age who have speech, language or learning disorders.",  
 "address": {  
 "street": "445 Sherman Avenue, Suite O",  
 "city": "Palo Alto",  
 "state": "CA",  
 "zip": "94306",  
 "type": "physical"  
 },  
 {  
 "attention": "Mid-Pen Speech & Language Clin",  
 "street": "445 Sherman Avenue, Suite O",  
 "city": "Palo Alto",  
 "state": "CA",  
 "zip": "94306",  
 "type": "mailing"  
 },  
 "service\_poc": [  
 "Kathy Kinkaid (kkinkaid@mpslc.com)"  
 ],  
 "hours": "By appointment",  
 "transportation": "SAMTRANS stops within 3 blocks, CALTRAIN within 3 blocks.",  
 "accessibility": [  
 "restroom",  
 "wheelchair"  
 ],  
 "languages": null,  
 "phones": [  
 {  
 "number": "650 321-8111",  
 "phone\_hours": "(Monday-Friday, 8-5)"  
 "type": "voice"  
 },  
 {  
 "number": "650 321-8122"  
 "type": "fax"  
 }  
 ],  
 "internet\_resource": [  
 "url": http://www.mpslc.com",  
 "email": "ddagenais@mpslc.com"  
 ],  
 "services": [  
 {  
 "audience": "Voice disorders, stutterers, children with developmental speech and/or language disorders",  
 "eligibility": null,  
 "fees": "Individual speech therapy $85 per half-hour session. Group therapy $60 per session. Fees for evaluation vary. Cash, checks and credit cards accepted.",  
 "how\_to\_apply": "Apply by phone.",  
 "service\_areas": [  
 "Alameda County",  
 "Contra Costa County",  
 "Marin County",  
 "San Francisco County",  
 "San Mateo County",  
 "Santa Clara County"  
 ],  
 "keywords": [  
 "DEVELOPMENTAL DISABILITIES SERVICES",  
 "Speech Therapy",  
 "HANDICAP SERVICES",  
 "Screening/Diagnosis",  
 "Speech and Language Evaluations"  
 ],  
 "wait": "Varies, usually not very long.",  
 "funding\_sources": [  
 "Fees"  
 ]  
 }  
 ]  
 }  
 ]  
 }  
]

The lowest common denominator for data exchange is the comma delimited text file or CSV. CSV are simple table structures. In order to represent complex data types such as arrays and objects, they need to be decomposed into individual tables that are joined to parent tables via a foreign key. The entity relationship diagram below shows how the CSV files are joined to form a single organization.



“Entity-relationship”

# Content

The data model is divided in organizations, locations and services. The following tables define the minimum set of data needed for a human services data specification.

## Organization

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Required | Detail |
| organization\_id | int | required | ID that uniquely identifies the organization |
| organization\_name | string | required | Name of the organization |
| FEIN | string | required | Federal Employe Identification Numner |
| AKA | string | optional | Alias or Also Known As |
| admin\_contact | array | optional | Administrative contact(s) |

## Locations

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Required | Detail |
| organization\_id | int | required | ID that uniquely identifies the organization |
| location\_id | int | required | ID that uniquely identifies the organization |
| location\_name | string | required | Name of the location |
| address | object | required | Physical address of location |
| description | string | required | Descrption of services provided at the location |
| hours | string | required | Hours of operation for the location |
| phone | array | required | Phone number for the location |
| mail\_address | object | optional | Mailing address of location |
| service\_contact | object | optional | Point of contact for service(s) provided |
| short\_desc | string | optional | Succinct description of services provided at the location |
| email | string | optional | Email address for location |
| fax | array | optional | Fax number for the location |
| tty | array | optional | TTY number for the location |
| languages | array | optional | Languages spoken at the location |
| transportation | string | optional | Public transportation alternatives near the location |
| accessibility | array | optional | Accessibility options available at the location |
| spatial\_location | object | optional | Coordinates of the location expressed as latitude and longitude using the WGS84 datum |

## Services

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Required | Detail |
| organization\_id | string | required | An ID that uniquely identifies an organization |
| location\_id | string | required | An ID that uniquely identifies a location |
| service\_id | string | required | An ID that uniquely identifies a service |
| how\_to\_apply | string | required | Description of how to apply to receive the service |
| eligibility | string | required | Requirements to receive the service |
| fees | string | required | Fees charged to receive the service |
| audience | string | optional | Clients that require the service |
| wait | string | optional | Wait times associated with the service |
| service\_areas | array | optional | Areas served |
| funding\_sources | array | optional | Source of funds used to support the service |

# Files

The following table list the files that are part of the Human Services Data Specification.

|  |  |  |
| --- | --- | --- |
| File | Required | Defines |
| organization.csv | required | Organization that provides human services. |
| location.csv | required | Locations that provide human services. |
| service.csv | required | Services provided at a location. |
| address.csv | required | location and mailing addresses |
| phone.csv | required | Phone numbers, hours, and types (voice, fax, tty). |
| keyword.csv | required | Key words for categorization. |
| service\_contact.csv | required | Service point of contact. |
| admin\_contact.csv | optional | Organization point of contact and phone number. |
| internet\_resource.csv | optional | Email addresses and urls. |
| service\_area.csv | optional | Area where service is available. |
| accessibility.csv | optional | Accessibility options available. |
| location\_coordinates.csv | optional | Coordinates for locations. |

## File Requirements

This section defines the format and contents of files.

* Data is saved in a comma-delimited file, string elements must be enclosed by double quotes
* The first line of each file must contain field names. Each subsection of the Field Definitions section corresponds to one of the files and lists the field names you may use in that file.
* Field names are case sensitive
* Field values may not contain tabs, carriage returns or new lines \*Field values that contain quotation marks or commas must be enclosed within quotation marks. In addition, each quotation mark in the field value must be preceded with a quotation mark. This is consistent with the manner in which Microsoft Excel outputs comma-delimited (CSV) files. For more information on the CSV file format, see http://tools.ietf.org/html/rfc4180.
  + The following example demonstrates how a field value would appear in a comma-delimited file:
  + Original field value: Contains “quotes”, commas and text
  + Field value in CSV file: “Contains ”“quotes”“, commas and text”
* Field values must not contain HTML tags, comments or escape sequences.
* Remove any extra spaces between fields or field names.
* Each line must end with a CRLF or LF linebreak character.
* Files should be encoded in UTF–8 to support all Unicode characters.

## Field Definitions

Mapping JSON objects to a flat file format requires normalization of data. This means breaking up objects and arrays into individual tables and adding keys to enable joining them in either a JSON object or in a relational database.

### Organizations

organization.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Details |
| organization\_id | string | required | Organization\_id is an ID that uniquely identifies an organization |
| organization\_name | string | required | The name of the organization |
| FEIN | string | required | Federal Employe Identification Numner |
| AKA | string | optional | Alias or Also Known As |

### Locations

location.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Details |
| organization\_id | string | required | An ID that uniquely identifies an organization |
| location\_id | string | required | An ID that uniquely identifies a location |
| location\_name | string | required | The name of the location |
| description | string | required | A verbose description of a location |
| short\_desc | string | optional | A succinct description of a location |
| hours | string | optional | The hours of operation of a location |
| transportation | string | optional | Availability of public transportation near a location |

### Service

service.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Details |
| organization\_id | string | required | An ID that uniquely identifies an organization |
| location\_id | string | required | An ID that uniquely identifies a location |
| service\_id | string | required | An ID that uniquely identifies a service |
| eligibility | string | required | Requirements to receive the service |
| fees | string | required | Fees charged to receive the service |
| how\_to\_apply | string | required | Description of how to apply to receive the service |
| audience | string | optional | Clients that require the service |
| wait | string | optional | Wait times associated with the service |

### Address

address.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| location\_id | string | required | An ID that uniquely identifies a location |
| type | string | required | Type can either be the physical address or the mailing address |
| street\_1 | string | required | Contains the street number, street prefix, street name, street postfix |
| city | string | required | The city or appropriate administrative unit where the location resides |
| state | string | required | The state or appropriate administrative unit where the location resides, If in the US, use the two letter abbreviation |
| postal\_code | string | required | The postal code where the location resides. |
| street\_2 | string | optional | Additional address information such as suite number |
| country | string | optional | The country where the location resides. Use [ISO 3166](http://www.iso.org/iso/home/standards/country_codes.htm) for country designators. |

### Phone

phone.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| location\_id | string | required | An ID that uniquely identifies a location |
| phone\_number | string | required | The phone number follows the pattern of area code, exchange prefix, line number; e.g. 5551234567 |
| phone\_type | string | required | Phone type values can be voice, fax, or tty. |
| phone\_extension | string | optional | Phone extension is prefixed with a lowercase “x” followed by the extension; e.g. x1111 |
| phone\_hours | string | optional | Hours that the phones will be ansewerd |
| country\_code | string | optional | Country telephone code prefix based in [ITU E.123](http://www.itu.int/rec/T-REC-E.123-200102-I/en) standard. |

### Administrative Contact

admin\_contact.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| organization\_id | string | required | An ID that uniquely identifies a location |
| contact\_name | string | required | Name of administrative contact |
| contact\_title | string | optional | Title of administrative contact |

### Service Contacts

service\_contact.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| location\_id | string | required | An ID that uniquely identifies a location |
| service\_poc | string | required | Name of service provider contact |
| poc\_title | string | optional | Title of point of contact |

### Internet Resource

internet\_resource.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| organization\_id | string | required | An ID that uniquely identifies a location |
| resource | string | optional | A URL or email address for contacting the organization |
| resource\_type | string | optional | Type of resource is either a URL in the form of http://organization.org/ or an email in the form of persont@organization.org |

### Accessibility

accessibility.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| location\_id | string | required | An ID that uniquely identifies a service |
| accomodation | string | required | A description of an accessibility service available at the location. E.g. wheelchair ramp, sign interpreter, etc. |

### Keyword

keyword.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| service\_id | string | required | An ID that uniquely identifies a service |
| keyword | string | required | A keyword that describes the service and can be used for classifying services under a taxonomy |

### Funding

funding.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| organization\_id | string | required | An ID that uniquely identifies a organization |
| source | string | required | A description of funding sources, e,g. county, state, private donors |

### Location Coordinates

location\_coordinates.csv

|  |  |  |  |
| --- | --- | --- | --- |
| HSD Name | Data Type | Required | Detail |
| location\_id | string | required | An ID that uniquely identifies a organization |
| latitude | number | required | The y coordinate expressed in decimal degrees in the WGS84 datum |
| longitude | number | required | The x coordinate expressed in decimal degrees in the WGS84 datum |

# Extension

The specification can be extended or modified with additional data elements. If the data type is simple it can be directly added to the organizations, locations and services table definitions. If the data type is complex, e.g. an array or an object, then an additional CSV file will be necessary. Foreign key should be assigned as needed with the intended purpose of the data.