# EmDigIt Data Schema

## Overview\_Table

This is the primary table for the database. The **rows** represent **entries,** defined as unique appearances of routes within a given edition. Entries consist of header data parsed into a) an origin location b) locations of intermediate postal stops and c) a destination location. The **columns** describe different **route characteristics** including: a) basic attributes (origin and destination) b) material properties (page numbers, encompassing edition) c) editorial descriptors (sum of posts, route attributes). Posts describe stops made by couriers and travelers to rest, acquire or distribute mail, and/or exchange horses. NB: routes may appear multiple times in the table, but each entry is considered a unique appearance and therefore row in the table.

|  |  |  |
| --- | --- | --- |
| Column Name | Description | Example |
| Entry\_ID | Assigns the unique identifier of an “entry.” An entry is determined by the edition’s original separation under a new heading, as signified by spacing or font change. | 476 |
| Edition\_ID | The unique identifier of the edition in which route may be found. Assigned in **Edition\_Table**. Usually consists of the author initials followed by year. | OC1608 |
| Date | The date of the edition in which the route is found, used for sorting purposes. Entries appear in order of date of publication. | 1577 |
| Route\_Description | The original text of the edition’s heading under which the route can be found. Brackets indicate historian’s best guess as to edition’s intent. Example: “[Nancy] a Metz” has no origin location provided in the heading, but a) falls under a similar route originating from Nancy b) falls under a larger subsection dealing with routes from Nancy or c) Nancy is the largest previously mentioned post in the vicinity of the first post stop along the route provided. | [Nancy] a Metz  Barcelon a miraflores monesterio carturos que es a una legua Burgos  Geneve a Chur par autre chemin |
| Source\_Location\_Name | The original text of the origin location as taken from the route heading, often preceded by preposition “from” (*da, von*) See exceptional cases above. | Milano |
| Source\_ID | The Location\_ID of the Source\_Location found by matching the Source\_Location to Location\_Name in the **Location\_Table.** | 45 |
| Target\_Location\_Names | The original text of the destination location as taken from the route heading. Multiple entries separated by “;”. Example: “Roma;Milano”. Individual target names in the following order: final destination, followed by intermediary stops in rough sequence. Example: Milano;Pisa;Genoa represents a route with a final destination of Milan, but intermediary locations of Pisa and then in Genoa. | Roma;Milano |
| Target\_ID\_All | The Location\_ID of the Target\_Location found by matching the Target\_Name to Location\_Name on the **Location\_Table.** ID’s follow the same ordering structure as the target name columns. Multiple entries separated by “;”. | 67;79;100 |
| Target\_ID\_0 | The Location\_ID of the last destination on a given route. | 67 |
| Route\_Comb | A representation of the route using the following format: “SourceID;TargetID[0];TargetID[n]” | 45;67;79;100 |
| Beg\_Pag\_# | The page number of the noted edition upon which the route begins. If edition is unpaginated, no number is given. “Verso” and “Retro” represented by whole number or decimal, such as “130” and “130.5.” | 167, 130.5 |
| End\_Pag\_# | The page number of the noted edition upon which the route ends. If edition is unpaginated, no number is given. “Verso” and “Retro” represented by whole number or decimal, such as “130” and “130.5.” | 122, 168.5 |

## 

## Route\_Table

This is a secondary table of the database that seeks to identify common routes across unique appearances. While an **entry** is defined as each appearance in an edition, a **route** can appear across editions. Each row represents a single common route as defined by the combination of its source location, target location, and any intermediary stops. The columns represent common route characteristics.

|  |  |  |
| --- | --- | --- |
| Column Name | Description | Example |
| Route\_ID | A unique ID assigned to each possible variation of source location, target location, and intermediary stops. | 1 |
| Route\_Comb | A representation of the route using the following format: “SourceID;TargetID[0];TargetID[n]” NOTE: Comb\_Route features a different sequence than Route\_Comb. | 57;189;200 |
| Dup\_Bw\_Titles | A route is defined as duplicated across editions if it appears at least twice and in at least two editions, as defined by the Edition\_Table. Boolean entries. | TRUE |
| Dup\_Bw\_Authors | A route is defined as duplicated across authors if it appears at least twice and in works by at least two authors, as defined by the Edition\_Table. Boolean entries. | FALSE |
| Source\_ID | The Location\_ID of the Source\_Location found by matching the Source\_Location to Location\_Name in the **Location\_Table.** | 6 |
| Target\_ID\_All | The Location\_ID of the Target\_Location found by matching the Target\_Name to Location\_Name on the **Location\_Table.** ID’s follow the same ordering structure as the target name columns. Multiple entries separated by “; | 133;185 |
| Target\_ID\_0 | The Location\_ID of the last destination on a given route. | 133 |
| Edition\_ID | The unique identifier of the edition in which route may be found. Assigned in **Edition\_Table**. Usually consists of the author initials followed by year. | OC1608 |
| Matched\_Entries | Matching Entry\_IDs from Overview\_Table, multi-value cells separated by ";" | 403;686;970;1860;2143;2796;8665;10818 |
| Last\_Date | Columns represent the date of publication of the last edition(s) in which route appears | 1665 |
| First\_Date | Columns represent the date of publication of the first edition(s) in which route appears | 1545 |
| Duration | Value equal to difference between last and first date of publications, as provided in preceding columns. Provides a rough duration of period of publication for a given route. | 120 |
| Source\_ID | The Location\_ID of the Source\_Location found by matching the Source\_Location to Location\_Name in the **Location\_Table.** | 1 |
| Source\_Location\_Name\_Standardized | A standardized name applied to Location\_Name and unique\_ID. | Bordeaux |
| Target\_ID | The Location\_ID of the Target\_ID\_0 found by matching the Target\_ID\_0 to Location\_Name in the **Location\_Table.** | 763 |
| Target\_Location\_Name\_Standardized\_Target | A standardized name applied to Location\_Name and unique\_ID. | Cape Finisterre |

## Edge\_Table

The Edge\_Table does not add additional information, but reformats the Route\_Table so that each row represents a relationship between two locations. Edges are generated to represent the Source to Destination, Source to Intermediary, and Intermediary to Destination.

|  |  |  |
| --- | --- | --- |
| Column Name | Description | Example |
| Edge\_ID | A unique ID assigned to each possible variation of source location and target location | 1 |
| Route\_ID | A unique ID assigned to each possible variation of source location, target location, and intermediary stops. | 1 |
| Source | The Location\_ID of the Source\_Location found by matching the Source\_Location to Location\_Name in the **Location\_Table.** | 6 |
| Target | The Location\_ID of the Target Location found by matching the Target Location to Location\_Name in the **Location\_Table.** | 6 |
| Seq\_Idx | An ID that identifies the Edge’s place in the Route sequence, starting from 0. A sequence ID of 0 corresponds to an Edge between the origin and the destination. Each subsequent edge between the origin and the intermediary locations has a value of 1 greater than the previous. | 0 |
| Comb\_Route | A representation of the route using the following format:  origin;dest, orig;int1;dest, and origin;int1;intn;dest. | 3;1027;70;161;973 |
| Edge\_Type\_Last\_Date | The latest date of publication in which the edge has been published | 1720 |
| Edge\_Type\_First\_Date | The earliest date of publication in which the edge appears | 1545 |
| Duration | Value equal to difference between last and first date of publications, as provided in preceding columns. Provides a rough duration of period of publication for a given edge. | 25 |

## Edition\_Table

This is a secondary table of the database providing further information on the postal itinerary editions. The **rows** represent **editions,** meaning a postal itinerary as differentiated by its constituent characteristics. The **columns** describe different **edition characteristics**, including author, year of publication, location of publication, etc.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Edition\_ID | Assigns a unique id to the edition. The format is the initials of the first name and surname of the primary author and the four digit year of publication. If the author is anonymous, the first two letters are “SA.” If no date is present, then the date is replaced by “SD.” If multiple editions are found sharing these characteristics but varying in others (such as publication location), these are denoted by an additional “A” or “B” at the end of the ID. | CE1552A |
| Edition\_Title | The original text of the short-title of the edition as used by short-title catalogues. | Itinerario delle poste |
| Edition\_Author | The full name of the author of the itinerary, as denoted by the title page. If no name provided, entry is “S.N.”. | Ottavio Codogno |
| Edition\_Location | The name of the location in which edition was produced, as denoted by the title page and standardized. If no name provided, entry is “S.L.” | Venice |
| Language | The original language of the current itinerary edition | Italian |
| Edition\_Publisher | The full name of the publisher of the itinerary, as denoted by the title page. If no name provided, entry is “S.N.”. Multiple publishes and/or name variants separated by semicolon. | Johann Cristoph Kißner;J.C. Kissner |
| Edition\_Date | The four digit year in which an edition was produced, as denoted by the title page. If no date in provided, the digits will be replaced by “SD”. | 1608 |
| Author\_VIAF | Virtual Internet Authority File (<http://viaf.org/>) ID provided where available. | VIAF116409342 |
| Author\_WikiData | WikiData (<https://www.wikidata.org>) ID provided where available. | Q16585233 |
| Title\_USTC | Universal Short Title Catalogue (<https://www.ustc.ac.uk/>) ID provided where available. | USTC 346874 |
| Title\_OCLC | Online Computer Library Center (<https://www.oclc.org>) ID provided where available. | 642943450 |
| Entered | Boolean value indicating whether route headers of indicated edition have been included in dataset | TRUE |

## Location\_Table

This is a secondary table of the database providing further geographic information on **locations**. The **rows** represent **locations,** meaning a place associable with a name within the itinerary and a set of coordinates. The **columns** describe different **location characteristics**, including a standardized name, region, and coordinates.

|  |  |  |
| --- | --- | --- |
| id | Assigns a unique id to each location. | 5 |
| Location\_Name\_Standardized | A standardized name applied to Location\_Name and unique\_ID. | Bordeaux |
| Location\_Name | Names found in original text that are believed by the historian to refer to a single location. Multiple entries are denoted by “;”. | Bordeauz; Bordeo; Bordeaulx |
| Geonames\_Name | The name of the location. |  |
| geonameId | Unique location ID for a given geonames location. | 3183364 |
| adminName1 | Name of administrative subdivision | Galicia |
| Region\_2 | Historically continuous political entity with which location can be associated, according to historian’s best guess. | Lombardy |
| Location\_Lat | Latitude for location | 43.9016 |
| Location\_Lng | Longitude for location | 1.89686 |
| Location\_LatLong | Latitude and longitude for location, as separated by a comma. | 46.1582121,1.8004451 |
| country\_code | The country code of the location | FR |
| country\_name | Modern nation state with which location can be associated, according to historian’s best guess. | Germany |
| Att\_Inf | Attribute determined by historian reflecting proximity to important geographic feature. | Mountain Pass |
| Att\_Inf\_Detail | Detail of attribute determined by historian reflecting proximity to important geographic feature. | Brenner Pass |
| Location\_Attribute\_Orig | String taken from the original text header indicating a qualitative evaluation of the route |  |
| Flag | True for any rows that do not contain the following country codes: DE, CZ, SK, PL, HU. | TRUE |
| Flag.note | String documenting researcher's reason for the application of a TRUE value in the "Flag" column |  |

## 