

**Virgin Trains West Coast**

**Staging Design specification**

Evolvi (Railblazers) data feed

Juanjo Diaz

Technical Consultant

**July 2018**

z

Contents

[1. Document Management 2](#_Toc520715902)

[2. Document Purpose 3](#_Toc520715903)

[2.1 In Document Scope 3](#_Toc520715904)

[2.2 Out of Document Scope 3](#_Toc520715905)

[3. High Level Overview 4](#_Toc520715906)

[3.1. Process Context Diagram 4](#_Toc520715907)

[3.2. Process Flow 5](#_Toc520715908)

[3.2.1 Adobe Campaign to SFTP 5](#_Toc520715909)

[3.2.2 SFTP to CRM 6](#_Toc520715910)

[4. Files 7](#_Toc520715911)

[4.1. Broadlog 7](#_Toc520715912)

[4.2. Delivery 7](#_Toc520715913)

[4.3. Tracking 8](#_Toc520715914)

[4.4. Extension 9](#_Toc520715915)

[5. Tables 12](#_Toc520715916)

[Migration.broad\_log 12](#_Toc520715917)

[Migration.Delivery 13](#_Toc520715918)

[Migration.tracking\_logs 14](#_Toc520715919)

[Migration.Extension 14](#_Toc520715920)

[6. Table Mapping 16](#_Toc520715921)

# Document Management

|  |  |  |  |
| --- | --- | --- | --- |
| **Version Control** | | | |
| **Version** | **Date** | **Author(s)** | **Sections Changed** |
| 0.1 | 31/07/2018 | Juanjo Diaz | Initial draft document |

| **Distribution List** | | |
| --- | --- | --- |
| **Organisation** | **Name** | **Role** |
| Merkle | Mark Jones | Client Lead / Project Manager |
| Merkle | Jez Cox | IBM Software SME |
| Merkle | Gary Newsome | Technical Consultant |
| Merkle | John Whittome | Business Analyst |
| Merkle | Peter Malherbe | AWS SME |
| Merkle | Mat Lynd | Development |
| Merkle | Juanjo Diaz | Technical consultant |
| Merkle | Steve Forster | Technical Solution Lead |
| Merkle | Avtar Aswell | Tester |

# Document Purpose

The purpose of this document is to describe the Evolvi data feed load process between Evolvi SFTP and Staging tables.

This document is intended for technical audience who is going to develop/maintain or test this process.

## 2.1 In Document Scope

The following items are in scope for this document:

* File definition
* Table definition
* Column mappings
* Data flow diagram

## 2.2 Out of Document Scope

The following items are out of scope for this document and will be covered to separate design documents:

* AWS connectivity

# High Level Overview

## 3.1. Process Context Diagram

Rail Blazers (Evolvi) will upload a file to VTWC SFTP on daily basis. Once the file is uploaded, SSIS will trigger [SFTP to S3 Lambda](https://teams.microsoft.com/_#/docx/viewer/teams/https%3A~2F~2Fmerkleinc.sharepoint.com~2Fsites~2FVirginTrainsWestCoast-CRMProject~2FShared%20Documents~2FGeneral~2FDevelopment%20Documentation~2FDesign%20Documentation~2FSFTP%20to%20S3~2FMerkle%20VTWC%20-%20SFTP%20to%20S3%20-%20v1.0.doc) function that will move that file to S3. Just after that, the process will continue and will load the file into a CRM PreProcessing table and then it will populate different tables on Staging that will complete the process and mark that file as processed.



Housekeeping process TBC.

## 3.2. Process Flow

The load process is split in 3 different steps.

**Step 1:** Evolvi uploads a file with all necessary information to VTWC SFTP.

**Step 2:** SSIS triggers a AWS Lambda function to download the file into S3 and load the data into PreProcessing schema.

**Step 3:** Load data from PreProcessing schema into Staging

Current document only describes steps 2 and 3. Step 1 is fully controlled by Evolvi.

## 3.2.1 Downloading file from SFTP to S3

For more information about this process, please read [SFTP to S3 documentation](https://teams.microsoft.com/_#/docx/viewer/teams/https%3A~2F~2Fmerkleinc.sharepoint.com~2Fsites~2FVirginTrainsWestCoast-CRMProject~2FShared%20Documents~2FGeneral~2FDevelopment%20Documentation~2FDesign%20Documentation~2FSFTP%20to%20S3~2FMerkle%20VTWC%20-%20SFTP%20to%20S3%20-%20v1.0.doc).

## 3.2.2 Loading data from file to PreProcessing

SSIS will load the content of the provided file into PreProcessing.Railblazers\_Data table. Mapping and data types are described on the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **FILE COLUMN\_NAME** | **TABLE COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** | **LENGTH** |
| Title | Title | NVARCHAR | TRUE | 255 |
| FirstName | FirstName | NVARCHAR | TRUE | 255 |
| Surname | Surname | NVARCHAR | TRUE | 255 |
| Company | Company | NVARCHAR | TRUE | 255 |
| AddressLine1 | AddressLine1 | NVARCHAR | TRUE | 255 |
| AddressLine2 | AddressLine2 | NVARCHAR | TRUE | 255 |
| TownCity | TownCity | NVARCHAR | TRUE | 255 |
| Postcode | Postcode | NVARCHAR | TRUE | 255 |
| ContactNumber | ContactNumber | NVARCHAR | TRUE | 255 |
| BusinessSector | BusinessSector | NVARCHAR | TRUE | 255 |
| BusinessStructure | BusinessStructure | NVARCHAR | TRUE | 255 |
| NumberOfEmployees | NumberOfEmployees | NVARCHAR | TRUE | 255 |
| Email | Email | NVARCHAR | TRUE | 255 |
| PromoCode | PromoCode | NVARCHAR | TRUE | 255 |
| Coast | Coast | NVARCHAR | TRUE | 255 |
| DateReceived | DateReceived | DATETIME | TRUE |  |
| MarketingOptIn | MarketingOptIn | BIT | TRUE |  |
| CreditOptIn | CreditOptIn | BIT | TRUE |  |
| DFTOptIn | DFTOptIn | BIT | TRUE |  |
| CreatedDateETL | N/A | DATETIME | FALSE |  |
| LastModifiedDateETL | N/A | DATETIME | FALSE |  |
| ProcessedInd | N/A | BIT | FALSE |  |
| DataImportDetailID | N/A | INT | FALSE |  |

## 3.2.3 PreProcessing to Staging Data Flow



## 3.2.3 PreProcessing to Staging process description

Once data is loaded into PreProcessing.Railblazers\_Data table, it potentially could be loaded into five different tables.

**Staging.STG\_Railblazers\_Evolvi**

This table is a copy of the PreProcessing table but designed to store historical data. It means that a housekeeping process should be applied [TBD].

**Staging.STG\_Individual**

We’ve to validate if the provided email or contact number (PreProcessing.Railblazers\_Data.Email, PreProcessing.Railblazers\_Data.ContactNumber) matches with any of the existing rows on Staging.STG\_ElectronicAddress table. As per Reference.AddressType table, we are using AddressTypeID = 3 for Email and AddressTypeID = 4 for Mobile.

If we can’t find a record for that email or contact number we will create a new row on Staging.STG\_Individual.

If we find a record, then we’ve to check if there is an IndividualID associated to that record and if so, look that CustomerID is empty. If that is the case, we’ll go to Staging.STG\_Individual table and we will update the data using data from PreProcessing.Railblazers\_Data table.

If from Staging.STG\_ElectronicAddress there is a CustomerID associate we won’t update Staging.STG\_Indiviual table, even if there is an IndividualID associated.

**Staging.STG\_Customer**

We won’t update information about customers using PreProcessing.Railblazers\_Data table data.

**Staging.STG\_CustomerPreferences**.

Based on PreProcessing.Railblazers\_Data.Email and PreProcessing.Railblazers\_Data.ContactNumber we will try to match this record against Staging.STG\_ElectronicAddress and get the CustomerID. If CustomerID is found, we will populate Staging.STG\_CustomerPreferences table using PreProcessing.Railblazers\_Data.MarketingOptin and PreProcessing.Railblazers\_Data.CreditOptin and PreProcessing.Railblazers\_Data.DFTOptin.

**Staging.STG\_IndividualPreferences.**

Based on PreProcessing.Railblazers\_Data.Email and PreProcessing.Railblazers\_Data.ContactNumber we will try to match this record against Staging.STG\_ElectronicAddress and get the IndividualID. If we cannot find a row or we find a row with IndividualID populated but CustomerID empty, we will populate Staging.STG\_IndividualPreferences table using PreProcessing.Railblazers\_Data.MarketingOptin and PreProcessing.Railblazers\_Data.CreditOptin and PreProcessing.Railblazers\_Data.DFTOptin

**Other consideration:**

These three preferences should be previously available on the Preference model. For more information about the preference model, please read [Preference Management document](https://teams.microsoft.com/_#/docx/viewer/teams/https%3A~2F~2Fmerkleinc.sharepoint.com~2Fsites~2FVirginTrainsWestCoast-CRMProject~2FShared%20Documents~2FGeneral~2FDevelopment%20Documentation~2FDesign%20Documentation~2FStaging~2FPreference%20Management~2FMerkle%20VTWC%20CRM%20-%20Pref).

# Files

Railblazers will produce a daily file that will be pushed to VTWC SFTP. This section contains information about the expected file format as well as VTWC SFTP file location.

## 4.1. File Name and Location

[TBD].

## 4.1. File Format

Table below contains the expected file format.

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **LENGTH** | **Description** |
| Title | 5 | Title of the prospect/customer for whom the booking has been done |
| FirstName | 25 | First name of the prospect/customer for whom the booking has been done |
| Surname | 25 | Surname of the prospect/customer for whom the booking has been done |
| Company | 150 | Company for what the prospect/customer is working for |
| AddressLine1 | 100 | Business Address Line 1 |
| AddressLine2 | 100 | Business Address Line 2 |
| TownCity | 50 | Business Town or City |
| Postcode | 10 | Business postal code |
| ContactNumber | 25 | Contact person of the prospect/customer who is going to travel |
| BusinessSector | 50 | Business Sector name |
| BusinessStructure | 50 | Finantial structure |
| NumberOfEmployees | 15 | Range number of number of employees |
| Email | 150 | Contact email of the prospect/customer who is going to travel |
| PromoCode | 35 | Promotion code used for that booking (if any) |
| Coast | 5 | Constant "West" |
| DateReceived | 16 | Date when the booking was done for that prospect/customer |
| MarketingOptIn | 5 | Flag to allow/deny marketing communications |
| CreditOptIn | 5 | Flag to allow/deny credit … |
| DFTOptIn | 5 | Flat to allow/deny data franchise transfer |

**END**