**Data Linkage Application**

The data linkage application is designed to disambiguate patients across sites / health systems, with minimal risk from transferring PHI by hashing the patient identifiers with a secured salt and hashing algorithm. Main Components of the application are:

1. Web based Project, User and Key Manager
2. Data Standardization, exceptions, and hashing
3. Disambiguation / matching

Each component is explained in more detail below:

1. **Web interface for Project, User and Key management**

This component is designed to manage project specific Salt distribution. It allows the project managers to create projects, add users and authenticate the users. The authenticated users upload their RSA public keys and download the encrypted Salt file.

1. **Data Standardization, exceptions, and hashing**

This component includes a data pipeline to digest and validate the data, standardize the data, manage data exceptions, create composite variables from patient identifiers and hash using SHA512 algorithm and Salt

1. **Disambiguation / matching**

This component allows the aggregator to merge files, disambiguate the hashes and assign Universal patient ID to matching patients

**Hashing and Matching Rules:**

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Composite identifiers (Hashed)** | **Record linkage logic** | **Matching rule** |
| 1 | First Name + Last Name + DOB (YYYY/MM/DD) (+ SSN) | match 1 against 1 and 2 | complete name and date of birth match (optional SSN) |
| 2 | Last Name + First Name + DOB (YYYY/MM/DD) (+ SSN) |  | transposed complete name and date of birth match (optional SSN) |
| 3 | First Name + Last Name + transposed DOB (YYYY/DD/MM) (+ SSN) | if no matches found above, match 1 against 3 | complete name and transposed date of birth's month and day match (optional SSN) |
| 4 | First 3 characters First Name + Last Name + DOB (YYYY/MM/DD) + SSN | if no matches found above, match 4 against 4 and 5 | partial name, date of birth match and SSN match |
| 5 | First 3 characters Last Name + First Name + DOB (YYYY/MM/DD) + SSN |  | transposed partial name, date of birth match and SSN match |
| 6 | First Name + Last Name + DOB (YYYY/DD/MM) + 1 day + SSN | if no matches found above, match 1, 2, 3 against 6 | complete name, date of birth plus 1 day and SSN match |
| 7 | First Name + Last Name + DOB (YYYY/DD/MM) + 1 year + SSN | if no matches found above, match 1, 2, 3 against 7 | complete name, date of birth plus 1 year and SSN match |