D&B Entity Matching

When Entity Matching to D&B Match Reference Data it is important to understand the impact of three key drivers of satisfaction.

All D&B core Entity Matching solutions are based on the same three key drivers of satisfaction.

- **Inquiry Quality**
  - It’s important to provide as much data as possible and to ensure that the data is as accurate as possible in the inquiry.
    - Name
    - Street Address
    - City
    - State/Province
    - Postal Code
    - Telephone Number
  - It’s important to use data fields for their intended use and to avoid extraneous data.
    - Try to ensure that a Name field has either a business name or a Contact Name and only that data.

D&B Proprietary & Confidential Information
• Extraneous data in a name field (e.g. “Do Not Use” or “likes to be contacted by telephone” etc) can impact the quality of the match result
  ▪ Try to ensure that the address field contains a street number and street name (depending upon the country)
  ▪ Mailing address is also a valuable match point
    • Extraneous data in an address field (e.g. C/O John Smith) can impact the quality of a match result.
  o Postal Codes are very predictive of a specific building in certain countries (e.g. UK, Netherlands etc) and can lead to a specific address.
  o A thin inquiry forces the Match solution into Search mode.
    ▪ An inquiry of Starbucks in New York City with no other address information is difficult to answer with a high degree of confidence as there are many Starbucks in New York City.

o D&B Match Reference Data
  o The D&B Match Reference data is updated with millions of changes every day
  o The granularity of the definition of a business could differ between D&B and the consumer of Entity Match services
    ▪ D&B may have a DUNS numbered entity for a Hospital, our customers may identify the flower shop and the pharmacy in the hospital as completely separate entities
    • Three separate inquiries for the hospital, the flower shop and the pharmacy could result in all three matching to the same Duns numbered entity if D&B does not recognize these as three separate legal entities.

o Match Environment
  o The D&B Entity Match environment is a key and algorithm based solution
    ▪ It is not a character based recognition solution
  o D&B Entity Match is non-deterministic
    ▪ The path of answering the inquiry is based upon the data available in the inquiry and the keys that are created from that data
      • Each inquiry will take its own path through the system based on the data provided
  o The goal of D&B Entity Matching is to find the one entity in the D&B database that most closely resembles the data provided in the inquiry
  o Cleanse, standardization and normalization processes are applied equally to inquiry data and candidate data to improve the opportunity to identify a high quality match
  o D&B provides Match Insight to help in the Data Stewardship process
    ▪ Confidence Code
    ▪ Match Grade String
    ▪ Match Data Profile
- Data Stewardship
  - D&B provides Match Insight as part of the answer to a match inquiry
    - Confidence Code
      - 10 – 4 – Tells the customer how confident they should feel about the answer provided. A Confidence Code 10 indicates a high degree of confidence. A Confidence Code 4 indicates a low degree of confidence

Confidence Code can be used to stratify results for auto-decisioning.

<table>
<thead>
<tr>
<th>Confidence Code</th>
<th>Quality of match improves</th>
<th>Match Grade pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>These patterns tend to have a strong match on name with decreasing correlation to individual address elements, supported by derived attributes.</td>
<td>Identifies why the customer should feel that way</td>
</tr>
<tr>
<td>9</td>
<td>These patterns tend to have more observed discontinuity in name or address (not both)</td>
<td>A pattern that represents each of the match points tells the customer which elements received a strong match and which received a less strong match</td>
</tr>
<tr>
<td>8</td>
<td>Strong match on name with poor match on address is a common theme (e.g. AFFF...)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Strong match on address with poor match on name is a common theme (e.g. FAAA...)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Internal use only (e.g., Intelligence Engine)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While there are millions of MatchGrade patterns that map up to the 10 Confidence Codes, it is helpful to think of stratification in terms of common themes.
Evaluation & Decisioning (E&D) determines and chooses the “best” candidate (batch) or candidates (on-line), and provides granular feedback to allow for complex decisioning actions.

- A qualitative evaluation is assigned for components and create a MatchGrade String (MGS)

<table>
<thead>
<tr>
<th>Business Name</th>
<th>Street Name</th>
<th>State</th>
<th>Telephone</th>
<th>Density</th>
<th>Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>A A A A B A Z B F A Z B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The MGS provides qualitative assessment between components of inquiry & reference info:
  - (A) Should be considered same: Smith Bros Clothiers vs. Smith Brothers Clothing
  - (B) Some similarity: Smith Bros Clothiers vs. Smith Clothing Store
  - (F) Should be considered not same: Smith Bros Clothiers vs. Smitty’s Bar and Pub
  - (Z) Missing information to conclude: One or both are blank or not used

- The MGS patterns are mapped to and used to rank-order matching candidates in context of Confidence Codes (CC) which range from 10 [highest] to 1 [lowest]

- Match Data Profile identifies to the customer which match points we used to make our decision
  - Which specific elements types did we match to (e.g. Primary Name, Tradestyle, Contact Name, Former Address etc).

- Using a combination of Confidence Code, Match Grade String and Match Data Profile can help a customer to identify the answers that they want to auto accept and the answers that they want to manually review or auto reject.