

Record Matching and De-duplication

One of the cornerstones of a patient registry implementation is the means to reliably match and link patient information collected from disparate sources. Typically, a **Master Patient Index**, or **MPI**, is created or acquired to meet this need. Various strategies can be employed within the MPI to match records or de-duplicate sets of records that appear to have the same patient represented more than once.

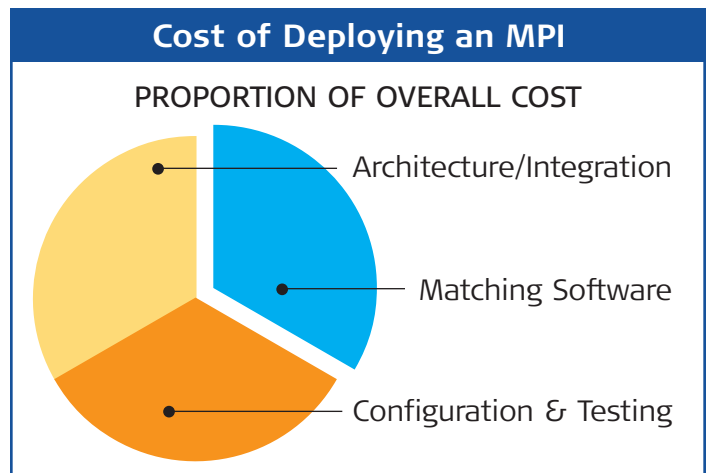
How do you select the right linkage and de-duplication strategy for your system's MPI?

There are two basic types of matching strategies:

- **Deterministic Matching** uses sets of predetermined rules to guide the matching process. The rules rely on a series of exact matches between data elements to identify when records match. It is most successful when the data is of relatively high quality or is dominated by reliable unique identifiers for records. Deterministic matching is less successful when the data is incomplete or inaccurate, when there are many spelling or transcription errors, or lots of inconsistencies (e.g., frequent name changes).

- **Probabilistic Matching** is a process whereby an estimate is made of the probability that two records are for the same person based on the degree to which certain fields in the two records match. Two thresholds are then set:
 - All record pairs whose probability is above the higher threshold are considered to be matches.
 - All record pairs whose probability is below the lower threshold are considered not to be matches.

The disposition of record pairs whose probability falls in between the two thresholds is considered to be



uncertain and they require additional review, likely by a trained staff member.

There are many software products available to perform either of these matching techniques, but software acquisition is only part of the puzzle. We estimate that only a third of the total cost of deploying an MPI pays for the software itself. Another third of the cost pays for the development of an appropriate architecture within which your MPI will function, as well as the integration of the MPI into the other components of your system. The final third pays for configuration and testing –by your staff or the MPI vendor–to ensure that the software works properly with your data and systems.



**A Collaborative
Approach to Health
Information Technology**

HLN Consulting, LLC provides a wide range of technology consulting services to healthcare. We focus especially on facilitating collaboration between the private and public health communities to support each other's missions. We design, develop, and support integrated public health systems with strong person matching capabilities. We are vendor and product neutral.

HLN is available to help in all aspects of registry conceptualization, planning, project management, and execution. Contact one of our specialists at info@hln.com to find out more information about how we can help *you* design a successful registry. Copyright ©2006 HLN Consulting, LLC

HLN Consulting, LLC
7072 Santa Fe Canyon Pl.
San Diego, CA 92129

Email: info@hln.com
Phone: 858-538-2220
Fax: 858-538-2209
<http://www.hln.com>