

Entity Resolution:

<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbncndYmRhdGF8Z3g6MzRjMmFjNzdhM2VIYzlhZA>

https://en.wikipedia.org/wiki/Jaro%E2%80%93Winkler_distance

Locality Sensitive Hashing:

<http://people.scs.carleton.ca/~maheshwa/courses/5703COMP/Talks/LSH/lsh.pdf>

Reservoir Sampling:

<http://dl.acm.org/citation.cfm?id=3165>

https://www.cs.duke.edu/courses/spring13/compsci590.2/slides/compsci590.02_spring13_lec1.pdf