Technical Challenges Along the Way to HIE Sustainability

May 20, 2014

Guest post by Egor Kobelev, software delivery manager — healthcare, [DataArt](#).

There are a lot of organizational and technical challenges health information exchanges (HIEs) struggle with while trying to deploy and maintain their platforms. One of the most complex organizational and administrative challenges is to achieve sustainability. While that is often an ultimate goal for HIEs, there is a huge amount of smaller technical challenges to meet, and the way those challenges are responded to often makes a difference for future HIE sustainability.

One of those typical tasks in the industry is a patient look up and mapping. There is a well-known issue when it comes to any sort of health data integration – the lack of a global unique patient identifier. Thousands of existing healthcare providers and payers use their own internal identifiers and there is no easy way to establish a relation between these. Social Security Numbers or similar national identifiers, while useful in some of scenarios, are not suitable for the purposes of healthcare record identification, primarily because of the risks of HIPAA rules violation.

The good part of the story is the amount of talks regarding a National Patient Identifier (NPI). For instance, HIMSS is proactively driving the initiative of introducing NPI, so that eventually patient mapping, which is currently a challenge, will be routine. However, the reality is that we are pretty far away from having NPI legislated and deployed in healthcare organizations nation-wide. At the same time, as many as 8 percent to 14 percent of patient records have errors caused by mismatching patient identifiers, which in turn causes hundreds of millions of dollars in spending to repair and reconcile the records. So, while we are waiting for NPI to come, what would be a solution which is HIPAA compliant, provides high accuracy, throughput, and minimizes manual interventions at the same time?

One of the most efficient approaches combines probabilistic machine matching, IDs mapping and the ability to fallback to a manual review when/if necessary. Let's say that healthcare organization A receives a message with patient health information from organization B. Both A and B maintain their own unique identification systems, which are incompatible. So A received patient health information, which, as a bare minimum, consists of name, date of

birth, gender, and organization B patient ID (B.ID for simplicity). Apparently the first step for A to take to match the patient is to go with probabilistic matching against its own patient database based on demographic information. Date of birth and gender are usually straightforward, name could be tricky, because of the complications caused by middle names, titles, suffixes, prefixes, etc. Those are often stored in unstructured formats, written differently and with no strict order. However, the majority of titles and prefixes are manageable by an exhaustive search against a dictionary, so that will leave just several combinations of first, middle and last names, handled either by strict equality or something like SoundEx, which works pretty well even in its canonical implementation.

Let's say organization A has managed to identify that the B.ID patient is actually the same person as A.ID. Now, this mapping between A.ID and B.ID needs to be stored on A's side, for future purposes. So the next time B will pass health information about the same patient, A could just lookup quickly against the mapping database to find out that B.ID is a patient A.ID, and this second time, there is no need for probabilistic matching.

While such an approach provides accurate patient matching in more than 99.99 percent of the cases, in some situations we cannot or do not want to rely fully on probabilistic matching. In such a case, it makes sense to append a detailed workflow with the ability to do a manual review of the mappings stored. Depending on the requirements, a review system like this could be as simple as just a list of all the mappings with ability to remove incorrect entries. Or something more advanced with the support of mappings sort order according to probability, batch operations or even automatic notifications of dubious mappings made.

One may argue that as long as such an approach is probabilistic, it still leaves room for errors and mistakes, which again leads to potentially incorrect care provisions and the necessity to repair and reconcile medical records. While such a concern is valid in the ideal world, let's admit that even having NPI in place, there is still a chance that staff or even a patient himself will make a typo while writing or typing his demographic information and NPI. So it all comes down to probability again. But no doubts, making NPI a part of the matching procedure, similar to the one described, will minimize the likelihood of mistakes and at the end of the day will improve the quality of healthcare services.