Big Data & Clinical Informatics

Client Overview

A leading clinical intelligence company that powers healthcare providers, life sciences and research organizations to make better-informed, more confident decisions by transforming unconnected data from multiple sources into real-world insights.

Problem Statement

Our client was trying to create the largest and most complete Electronic Health Record (EHR) database with a longitudinal view of individual patients and patient populations by gathering, normalizing and analyzing data from disparate IT systems — including EHRs, Practice Management Systems and claims data.

The aim was to utilize the valuable yet unstructured clinical data stored in various EHR warehouses by transforming it into consistent structured data for predictive analysis, modeling and interpretation of health dynamics.

The main challenge was to figure out how to handle data from disparate sources and normalize them in a standard format which can - be eventually used to generate analytics and make predictive decisions. The different data formats followed by various EHR systems posed a unique challenge of data interpretation and employing automated means to interpret the data correctly was a challenge in itself. Issues related with data provenance, normalization of the data & validations of the data correctness were the main concerns in this whole process.

The client was looking for a technology partner who had extensive experience in clinical informatics that could be leveraged to address the challenges faced by client. The client was looking specifically for a technology partner who had prior experience and deep understanding of EHRs, data warehousing, analytics as well as experts in content creation.

Solution

OPTRA team engaged with the client's informatics team and understood the challenges. OPTRA wrote the data acquisition adapters to read data from various EHR systems (CERNER, Epic just to name a few). Using the ready building code blocks, OPTRA normalized the data to client specified standard format. This data was then validated for data accuracy using automated as well as manual routines/SOPs.

OPTRA worked on a three pronged approach of applying big data analytics to EHR data and utilizing the rich data to:

- Build a "clinical decision support system," that healthcare providers can use to evaluate their proposed treatments and practice evidence-based medicine
- Create a platform for integrating real time patient-level data to reduce clinical trial cycle times, predict clinical trial outcomes early and accelerate drug discovery process
- Aggregate patient data from longitudinal medical records to yield improved insight into a drug's efficacy and risk-benefit profile

The following points were some of the key considerations while delivering the solution:

- **Data acquisition & recording**
  - Designed metadata acquisition systems to generate suitable metadata and data systems that carry
    - the provenance of data and its metadata through data analysis pipelines.
- **Data Analysis:**
  - Eyeball the native format data
  - Review tables of interest to define links
  - Identify extraction or loading problems such as coded fields with text, noncodes or transposition of columns
- **Data Provenance:** To identify the location in local information systems of the data elements and values of
  - interest to the client
  - Identify data elements of interest
  - Note all locations where elements are found or expected
  - Reconcile multiple locations into a single provenance statement which may include conditional or inclusive cascades
- **Normalization:** Transformations of all data types applied to local data to conform to the

client's standard

- Review or create client's standard representation
- Work with domain scientists to create effective and intelligent database designs
- Map controlled vocabulary, export to database
- NLP evaluation of long text values using a knowledge base constructed from related data for semantic interpretation and intelligent imputation based on relatedness
- Confirm interpretation of numeric values
- Document normalization specifications

✓ **Validation:** Specification of the qualifiers applied to normalized local data to ensure it conforms to the quality standards and intentions of the client

- Structural validation
- Semantic validation
- Referential validation

✓ **Verification:**

- Automated Manual verification and annotation

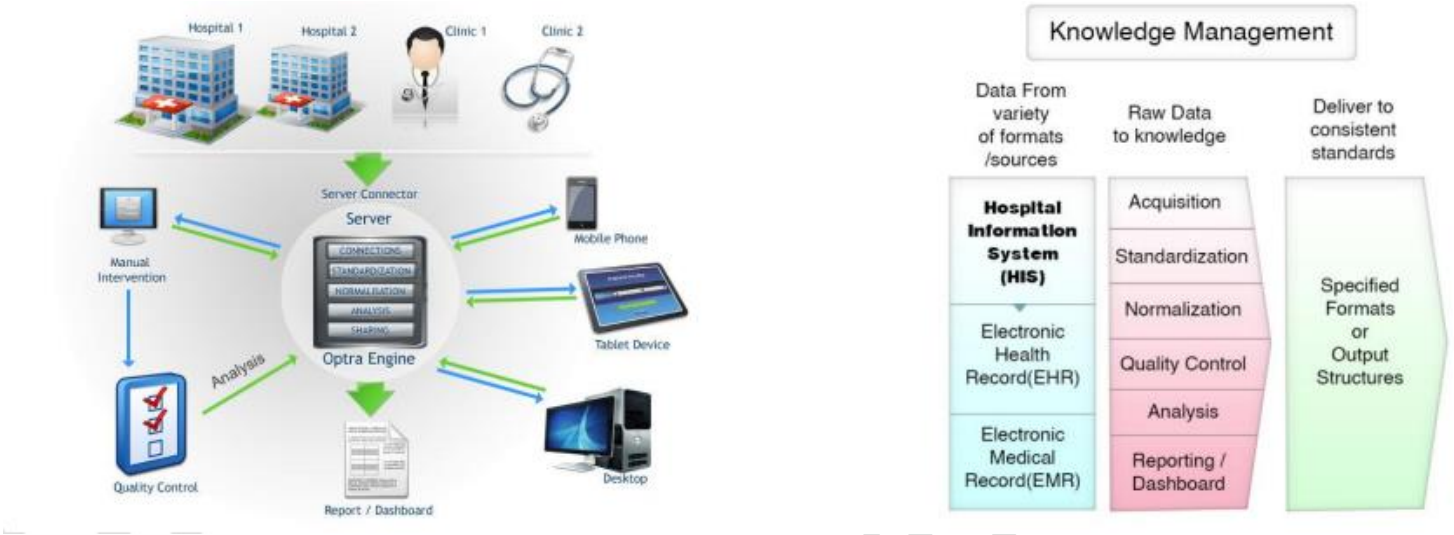
Once the data was generated, OPTRA wrote algorithms (using neural networks, Artificial Intelligence etc.) to generate analytics based on the data. Various predictive algorithms were written to search patient information, treatment information etc. which was later extended to assist healthcare companies and insurance companies for patients health history.

OPTRA used Big Data analytics (using Hadoop, MapReduce, ETLs etc.) to create scalable, efficient and automated data pipelines capable of handling petabytes of data. OPTRA's content creation and verification team was very instrumental in ensuring that the data generated was checked for quality as well as correctness. The key job was to ensure the data was mapped correctly from hospital EHR biomedical terminologies to client's standard mapping tables.

The following are some of the implementations, classifications and references used during the project:

- ICD-9-CM for diagnoses and in-patient procedures
- ICD-10-CM/PSC and GEMs
- CPT4 for ambulatory procedures and visits
- NDC codes for individual medications
- AHRQ CCS categories for condition roll-ups
- First Data Bank F-codes (For allergies)
- LOINC and SNOMED (supported)

The following diagram depicts the solution that was proposed and built for the client. Of course, there were client specific rules embedded in the final solution.



By leveraging ready building blocks, a wealth of experience and excellent operational efficiency, Optra was able to deliver on the client's goals.

Benefits

- The aggregation of individual EHRs and its analysis by algorithms-based clinical decision making through evidence-based medicine
- Big data analytics efficiently mined the clinical data warehouses filled with valuable unstructured data to help doctors make decisions about patient treatment.
- Central repository of the patient data with clinical informatics to identify at-risk patients (predictive)
- Metrics now an integral ingredient of the system
- Better management of patients and those with chronic conditions
- Improved clinical outcomes from systematic and proactive patient follow-ups
- Opportunities for improvement and standardization identified
- Support for big data analytics and cloud deployment

Optra Systems, Inc.
530 Lakeside Drive, Ste 250, Sunnyvale, CA 94085
Tel: +1-408-524-5300, Fax: +1-408-524-5302, Email: info@optrasystems.com

Technology Environment

- Big Data analytics (Hadoop & MapReduce)
- Oracle database, ETLs
- Semantics, HL7
- Predictive analysis algorithms in Java/C++
- Java/J2EE, Spring, Struts2, Maven, JQuery, Web services
- RIA development with JavaScript/Flex
- Visual studio 2010.

About Optra Systems

Optra Systems is an ISO-certified global organization with deep domain expertise in medical devices, lab automation, life science informatics and healthcare IT solutions. The company provides a fully-scalable, cost-effective OptiShore™ delivery model. This enables customers to choose the optimal balance between on-site, on-shore, and off-shore development that will best address their budget and collaboration requirements. With Optra Systems, customers are able to shrink their time-to-market by leveraging practical, building-block based solutions. Committed to clear communication and total transparency, the company consistently meets or exceeds its clients' expectations. Offering a full complement of expert engineering and consulting services, Optra Systems is aligned to real business needs applied over the entire product development lifecycle. The robust, scalable and efficient IT infrastructure of the company, together with its outstanding project management team, consistently ensures superior results. Optra Systems 's global delivery model helps its customers cut costs by about 50% without compromising on quality and realize a 200% improved production cycle.

Visit Optra Systems today: http://www.optrasystems.com

Contact Optra Systems Today

Optra Systems, Inc.
530 Lakeside Drive, Ste 250
Sunnyvale, CA 94085
Tel: +1-408-524-5300
Fax: +1-408-524-5302
Email: info@optrasystems.com