

Preliminary Study of EMR coding of billing project

Jary Du

Introduction

Background

Medical coding and billing is the process of assigning codes to the documentation of a patient's illnesses and treatments, and then billing insurance for reimbursement. Due to the complexity of the coding and billing process, there are many opportunities existing for problems. The American Medical Association blames errors in claims last year for an extra \$17 billion in administrative costs to physicians. Nearly 20 percent of the claims doctors get back from insurance companies have errors, according to the group. (Tribble, 2012) In addition to existing problems of medical coding process, the Department of Health and Human Services announced the new ICD-10 implementation is scheduled for October 1, 2014. Due to the significantly increased number of codes, the change in the number of characters per code, and increased code specificity, the moving to ICD-10 is expected to impact all physicians and creating massive amount of problems and also opportunities. At the same time, in the other side of the pacific ocean, China is having a medical reform which is impacting every aspect of Chinese healthcare including medical coding and billing system. Due to its drastic changes in healthcare and much higher volume of patients, Chinese medical coding system is facing challenges and also providing opportunities in this field.

Purpose of this project

Considering the existing problems in medical coding and the upcoming transition from ICD-9 to ICD-10, this project starts with exploring problems in current medical coding and billing process, and looking into the significant challenges that physicians are facing in the conversion from ICD-9 to ICD-10. This project also analyzes the current market of medical coding business and the incumbents in this market. With the comparison of current players, this project proposes a new approach to tackle the discovered problems by automating the whole medical coding process with natural language processing and question answering technologies. In addition to U.S medical coding market, this project also analyzes Chinese medical coding system and market with considering the upcoming future due to the Chinese medical reform. The current landscape of Chinese natural language processing technology and the opportunities it brings in the Chinese medical coding field are also the interest of this project.

Table of Contents

Introduction	1
<i>Background</i>	1
<i>Purpose of this project</i>	1
Table of Contents	2
Medical coding	3
<i>Current medical coding and billing process:</i>	3
Glossary of terminology in EMR coding procedure	4
<i>CPT</i>	4
<i>ICD-9</i>	4
<i>ICD-10</i>	4
<i>SNOMED CT</i>	4
<i>General equivalent mapping:</i>	5
Current problems	5
Proposed solution	6
<i>Automate the medical coding process:</i>	6
<i>Analysis engines design</i>	7
Type system design	8
<i>Study of ICD-9-CM coding guidelines</i>	8
<i>Study of CPT coding guidelines</i>	9
<i>Study of SNOMED clinical term guidelines</i>	10
Building pipeline	11
<i>Type system</i>	11
Competitor analysis	13
<i>ProVation MD coding engine</i>	13
<i>AccelaSMART</i>	13
<i>PrecisionCode</i>	13
<i>Product feature comparison table</i>	14
Accountable care act	14
<i>Problem</i>	14
<i>Accountable care organizations</i>	15
Medical coding in China	16
<i>Medical reform in China</i>	16
Natural language processing in Chinese	17
<i>Stanford NLP research in Chinese</i>	17
<i>Chinese word segmentation</i>	17
<i>Parsing and grammatical relations</i>	17
<i>Tsinghua University NLP</i>	17
<i>Companies involved in Chinese NLP</i>	18
<i>Opportunities</i>	19
Reference	Error! Bookmark not defined.

Medical coding

The medical billing process is an interaction between a health care provider and the insurance company (payer). The entirety of this interaction is known as the *billing cycle* sometimes referred to as Revenue Cycle Management. This can take anywhere from several days to several months to complete, and require several interactions before a resolution is reached. The relationship between a health care provider and insurance company is that of a vendor to a subcontractor. Health providers are contracted with insurance companies to provide health care services. (American Academy of Professional Coders, 2012)

Current medical coding and billing process:

The interaction begins with the office visit: a doctor or their staff will typically create or update the patient's medical record. This record contains a summary of treatment and demographic information including, but not limited to, the patient's name, address, social security number, home telephone number, work telephone number and their insurance policy identity number.

Upon the first visit, the provider will usually give the patient one or more diagnoses in order to better coordinate and streamline their care. In the absence of a definitive diagnosis, the reason for the visit will be cited for the purpose of claims filing. The patient record contains highly personal information, including the nature of the illness, examination details, medication lists, diagnoses, and suggested treatment.

The extent of the physical examination, the complexity of the medical decision-making and the background information (history) obtained from the patient are evaluated to determine the correct level of service that will be used to bill the insurance.

The level of service, once determined by qualified staff is translated into a standardized five-digit procedure code drawn from the Current Procedural Terminology database.

Once the procedure and diagnosis codes are determined, the medical biller will transmit the claim to the insurance company (payer). (G&G Advanced Medical Consulting, Inc., 2012)

Glossary of terminology in EMR coding procedure

CPT

Current Procedural Terminology code set describes medical, surgical, and diagnostic service and is designed to communicate uniform information about medical services and procedures among physicians, coders, patients, accreditation organizations, and payers for administrative, financial, and analytical purposes.

New editions are releases each October. The current version is the CPT 2011. It is available in both a standard edition and a professional edition.

CPT coding is similar to ICD-9 and ICD-10 coding, except that it identifies the services rendered rather than the diagnosis on the claim. (American Medical Association, 2012)

ICD-9

International Classification of Diseases, 9th revision

It is known as a health care classification system that provides codes to classify diseases and a wide variety of signs, symptoms, abnormal findings, complaints, social circumstances, and external causes of injury or disease. Under this system, every health condition can be assigned to a unique category and given a code, up to six characters long. Such categories can include a set of similar diseases. (Torrey, 2012)

ICD-10

International Classification of Diseases, 10th revision

The code set allows more than 14,400 different codes and permits the tracking of many new diagnoses. Using optional subclassifications, the codes can be expanded to over 16,000 codes. Using codes that are meant to be reported in a separate data field, the level of detail that is reported by ICD can be further increased, using a simplified multi-axial approach. HHS has announced the final rule that delays the ICD-10 compliance date from October 1, 2013 to October 1, 2014. (Torrey, 2012)

SNOMED CT

Systematized Nomenclature of Medicine – Clinical Terms, is a systematically organized computer processable collection of medical terms providing codes, terms, synonyms and definitions covering diseases, findings, procedures, microorganisms, substances, etc.

It allows a consistent way to index, store, retrieve, and aggregate clinical data across specialists and sites of care. It also helps in organizing the content of medical records, reducing the variability in the way data is captured, encoded and used for clinical care of patients and research.

It is a structured collection of medical terms that are used internationally for recording clinical information and are coded in order to be computer processable. It

covers areas such as diseases, symptoms, operations, treatments, devices and drugs. Its purpose is to consistently index, store, retrieve, and aggregate clinical data across specialties and sites of care. It helps organizing the content of electronic health records systems, reducing the variability in the way data is captured, encoded and used for clinical care of patients and research. (American Medical Association, 2012)

General equivalent mapping:

Centers for Disease Control and Prevention created the national version of the General Equivalence Mapping to ensure the consistency in national data is maintained.

The GEMs are a tool to assist with converting larger International Classification of Disease, 9th Edition, Clinical Modification (ICD-9-CM) databases to ICD-10-CM and ICD-10-PCS. (Torrey, 2012)

Current problems

Insurance claim forms can be complex to complete, often resulting in incorrect information. Incorrect information like patient demographics and invalid insurance company codes are common reasons for claim denial.

Incorrect code assignment is caused by a misunderstanding of the diagnosis or procedure performed. Incorrect diagnosis or procedure code assignment causes problems for reporting, statistics and reimbursement. Insufficient or illegible clinician documentation can result in incorrect code assignment and billing errors. Codes assigned without supportive documentation may result in insurance payment denial.

Current medical coding and billing process is majorly performed by human labor and sometime make it very labor-intensive and time-consuming. Considering the claim denial caused by error coding, this process tends to be much longer than it should be.

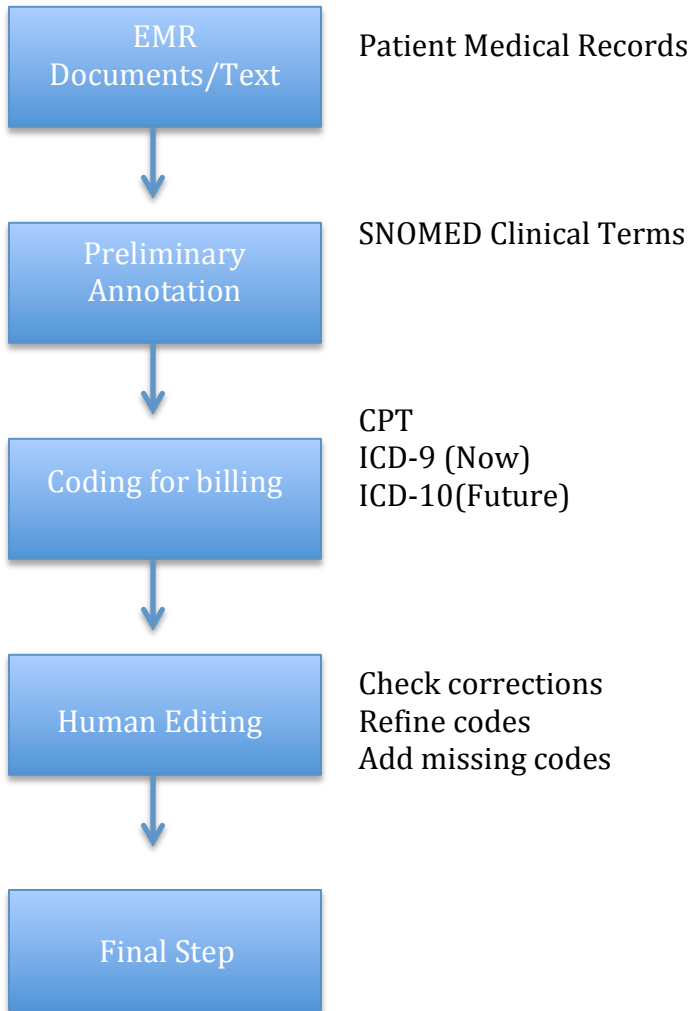
Switching to ICD-10 however, is a massive overhaul of the coding scheme and will require field size expansion, change to alphanumeric composition, and complete redefinition of code values and their interpretation. It will impact providers, payers, software vendors, clearinghouses, laboratories, insured's and potentially many other entities, especially medical coders.

How to use software technology to improve medical coding process is the primary goal of this preliminary study.

Proposed solution

Automate the medical coding process:

Logical architecture of Electronic Medical Record Coding



Type system design

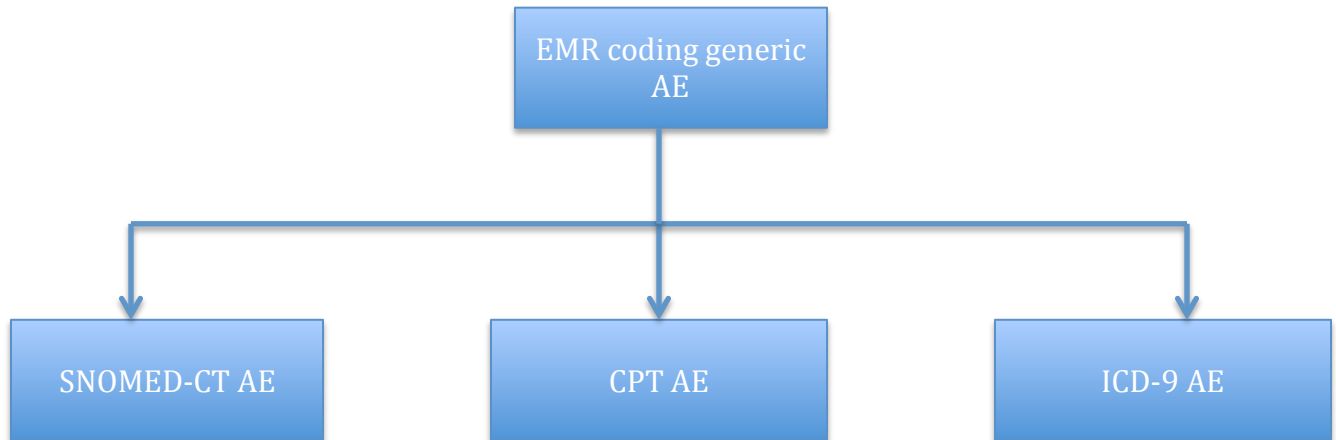
- SNOMED-CT type design
- CPT type design
- ICD-9 type design

Common analysis system design

- Break down the whole document into segments
- Start with heading and body text

Analysis engines design

- EMR coding generic analysis engine (annotator)
- SNOMED-CT analysis engine (annotator)
- CPT analysis engine (annotator)
- ICD-9 analysis engine (annotator)



Integration and Demonstration as a pipeline

- Setup the whole pipeline

'Gold-standard data' evaluation

- Get sample data
- Label sample data
- Evaluation with CSE framework
- Create metrics

Type system design

Study of ICD-9-CM coding guidelines

These guidelines have been developed to assist both the healthcare provider and the coder in identifying those diagnoses and procedures that are to be reported. The entire medical record should be reviewed to determine the specific reason for the encounter and the conditions treated. (American Medical Association, 2008)

General coding guidelines

Coder should use of both alphabetic index and tabular list when locating and assigning a code. Coder needs locate each term in the Alphabetic Index and verify the code selected in the Tabular List. Diagnosis and procedure codes are to be used at their highest number of digits available.

ICD-9-CM diagnosis codes are composed of codes with either 3, 4, or 5 digits. Codes with three digits are included in ICD-9-CM as the heading of a category of codes that may be further subdivided by the use of fourth and/or fifth digits, which provide greater detail. A three-digit code is to be used only if it is not further subdivided. When fourth-digit subcategories and/or fifth-digit sub-classifications are provided, they must be assigned. A code is invalid if it has not been coded to the full number of digits required for that code.

ICD-9-CM Volume 3 procedure codes are composed of codes with either 3 or 4 digits. Codes with two digits are included in ICD-9-CM as the heading of a category of codes that may be further subdivided by the use of third and/or fourth digits, which provide greater details. The appropriate code or codes from 001.0 through V89.09 must be used to identify diagnoses, symptoms, conditions, problems, complaints or other reasons for the encounter/visit. The selection of codes 001.0 through 999.9 will frequently be used to describe the reason for the admission/encounter. Those codes are from the section of ICD-9-CM for the classification of diseases and injuries like infectious and parasitic diseases; neoplasms; symptoms, signs, and ill-defined conditions, etc. (American Medical Association, 2008)

Type system design of ICD-9-CM

Code (string)	053.14	E800	V30.01
Code Type (character)		E	V
Main Category Code (numeric)	053	800	30
Subdivided Category-I Code (numeric)	1		0
Subdivided Category-II Code (numeric)	4		1
Description (string)	Herpes zoster myelitis	Accidental poisoning by tranquilizers	Single liveborn, born in hospital, delivered by cesarean section

Study of CPT coding guidelines

Current Procedural Terminology is a coding system to convert widely accepted, uniform descriptions of medical, surgical and diagnostic services rendered by healthcare providers into five-numeric codes. (Rouse, 2010)

There are three categories of codes in CPT.

Category I codes is the major part of CPT4 2009 manual which contain list of codes which begin with code 00100 and ends with code 99607. Category I codes are also called as the regular CPT codes. Category I CPT codes are the codes used to report the physician's services. All the category I codes are numeric. They are divided into six sections: Evaluation & Management services, Anesthesia, Surgery, Radiology, Pathology & Laboratory & Medicine.

CPT manual contains category II codes which are meant to measure the performance of certain services. Use of category II codes is not mandatory for coding purpose.

CPT category III codes are a list of codes for new and emerging technology, services and procedures. The codes representing the services listed in category III may be incorporated into regular CPT codes in future if approved by FDA. (Spinello, 2009)

Type system design of CPT codes

Code (string)	01990	1220F	0053T
Type (character)		F	T
Main Category (string)	Category I	Category II	Category III
Subdivided Category-I Code (numeric)	1	1	0
Subdivided Category-II Code (numeric)	990	220	053
Description (string)	Physiological support for harvesting of organ(s) from brain-dead patient	Patient screened for depression (SUD)	Replacement or repair of implantable component or components of total replacement heart system (artificial heart), excluding thoracic unit

Study of SNOMED clinical term guidelines

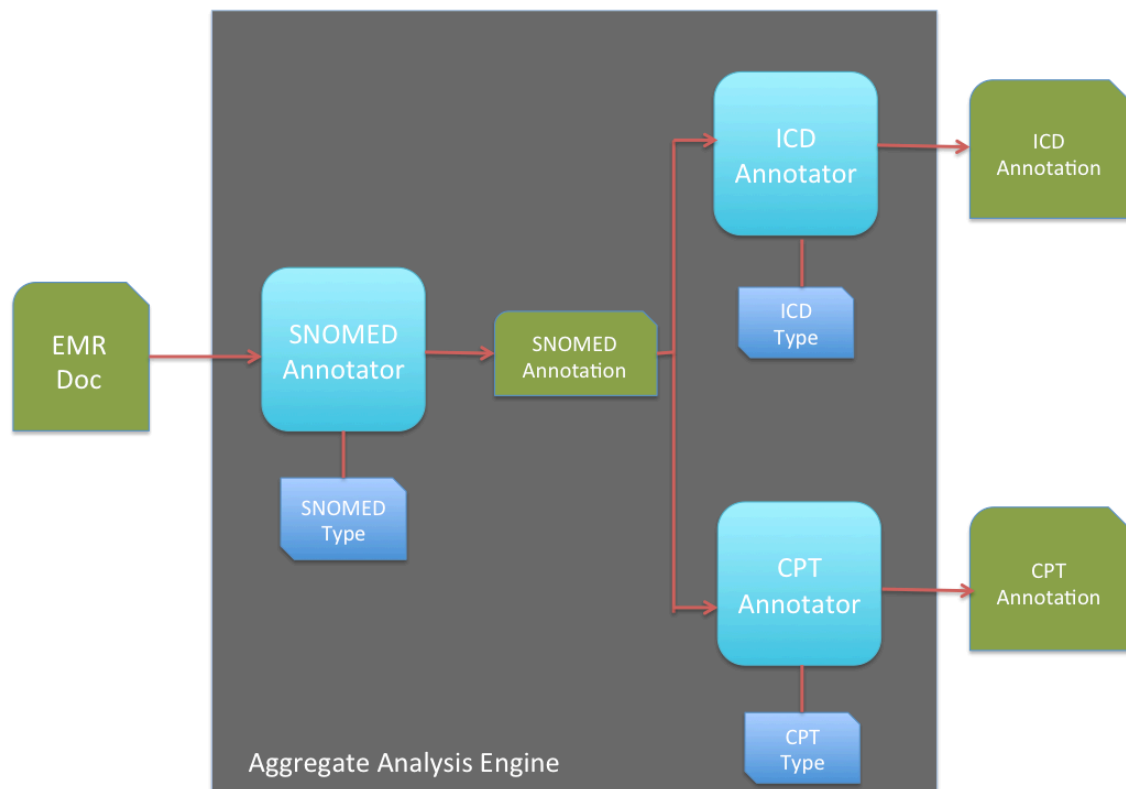
Systematized Nomenclature of Medicine—Clinical Terms is a comprehensive clinical terminology. SNOMED CT is one of a suite of designated standards for use in U.S. Federal Government systems for the electronic exchange of clinical health information and is also required standard in interoperability specification of the U.S. Healthcare Information Technology Standard Panel.

SNOMED Clinical Terms is a comprehensive clinical terminology that provides clinical content and expressivity for clinical documentation and reporting. It can be used to code, retrieve, and analyze clinical data. (International Health Terminology Standards Development Organization, 2012)

Type system design of SNOMED CT codes

SNOMED ID (string)	D3-15134	DB-61070
Concept ID (numeric)	15990001	2751001
Fully Specified Name (string)	Acute myocardial infarction of posterolateral wall (disorder)	Fibrocalculous pancreatic diabetes (disorder)
Preferred Term Description ID (numeric)	27071012	5649018
Preferred Term (string)	Acute myocardial infarction of posterolateral wall	Fibrocalculous pancreatic diseases
Synonyms (array of synonym)		

Building pipeline



Type system

CPT

Type System Definition

▼ Types (or Classes)

The following types (classes) are defined in this analysis engine descriptor. The grayed out items are imported or merged from other descriptors, and cannot be edited here. (To edit them, edit their source files).

Type Name or Feature Name	SuperType or Range	Element Type	
<input type="checkbox"/> edu.cmu.lti.oacb.CPT	uima.tcas.Annotation		Add Type
code	uima.cas.String		Add...
code_type	uima.cas.String		Edit...
main_category_code	uima.cas.Integer		Remove
sub_category_I	uima.cas.Integer		Export...
sub_category_II	uima.cas.Integer		
description	uima.cas.String		JCasGen

SNOMED

Type System Definition

▼ Types (or Classes)

The following types (classes) are defined in this analysis engine descriptor.
The grayed out items are imported or merged from other descriptors, and cannot be edited here.
(To edit them, edit their source files).

Type Name or Feature Name	SuperType or Range	Element Type	
<input type="checkbox"/> edu.cmu.lti.oacb.SNOMED	uima.tcas.Annotation		Add Type
id	uima.cas.String		Add...
concept_id	uima.cas.Integer		Edit...
fully_specified_name	uima.cas.String		Remove
preferred_term_id	uima.cas.Integer		Export...
preferred_term_description	uima.cas.String		
synonyms	uima.cas.StringArray	↔	JCasGen

ICD

Type System Definition

▼ Types (or Classes)

The following types (classes) are defined in this analysis engine descriptor.
The grayed out items are imported or merged from other descriptors, and cannot be edited here.
(To edit them, edit their source files).

Type Name or Feature Name	SuperType or Range	Element Type	
<input type="checkbox"/> edu.cmu.lti.oacb.ICD	uima.tcas.Annotation		Add Type
code	uima.cas.String		Add...
code_type	uima.cas.String		Edit...
main_category_code	uima.cas.Integer		Remove
sub_category_I	uima.cas.Integer		Export...
sub_category_II	uima.cas.Integer		
description	uima.cas.String		JCasGen
<input type="checkbox"/> edu.cmu.lti.oacb.SNOMED	uima.tcas.Annotation		
id	uima.cas.String		
concept_id	uima.cas.Integer		

Competitor analysis

ProVation MD coding engine

The flexible ProVation MD coding engine which automatically generate CPT and ICD codes based on physician documentation, easily accommodates the expanded ICD-10 code set and can function effectively in a dual coding environment.

ProVation MD allows physicians to quickly generate complete procedure notes at the point of care. As the physician makes menu selection, the software's powerful coding engine generates the associated CPT and ICD codes, as well as CCI edits. The physician then reviews and accepts the codes. The resulting procedure note is ready for coder review and same day billing. This drives revenue recovery because physicians and facilities are actually reimbursed for the procedures they perform – with no worries about accidental overcoding or defensive undercoding.

AccelaSMART

AccelaSMART™ – White Plume's intelligent, customizable rules-based management and automated workflow engine for code editing, front-end denial prevention, exceptions, variations, change processing and MD/practice-specific preferences.

AccelaSMART enables your billing staff to access the same technology payers use to deny your claims. Every charge is reviewed against a comprehensive code edit, custom rule and workflow engine. Errors or missed charges are automatically routed to the practice administrative staff along with real-time suggestions to correct each encounter. Corrected clean encounters are then uploaded into the practice management system for immediate processing.

The AccelaSMART resolution engine makes it easy to automate special circumstances, practice preferences and even physician quirks. No other system can provide the same level of automation and adjustments for keeping practice throughput optimized. (White Plume, 2012)

PrecisionCode

ePrecisionCode® is a fully web-based code scrubbing and comprehensive claims editing software, available as either a web-based or desktop solution. It provides a complete coding and compliance solution for practices of any size or specialty. ePrecisionCode is NOT just a simple look-up or reference tool, but a powerful true state-of-the-art application that captures all earned charges in a user-friendly, timely manner and helps accelerate reimbursement. (PrecisionCode, 2012)

Product feature comparison table

	Automatically generate CPT/ICD codes	ICD/CPT lookup table	Charges are reviewed	Streamline process	Identify possible omissions	Present code errors	Provide suggestions for errors	Support ICD-10 transition
Provation MD	★			★				★
UPMC TDC	★			★	★			
AccelaSmart			★	★	★	★	★	
Alpha II - CodingSystem		★	★			★		
PrecisionCode		★						
NueMD		★				★		
Flash Code		★						
TruCode		★		★				★

Accountable care act

The Affordable Care Act includes a number of policies to help physicians, hospitals, and other caregivers improve the safety and quality of patient care and make health care more affordable. By focusing on the needs of patients and linking payments to outcomes, these delivery system reforms will help improve the health of individuals and communities and slow cost growth. (Health & Human Services Press Office, 2011)

Problem

One in four Americans – and two out of three Americans over 65 – have multiple chronic conditions. According to a new report issued by the Centers for Medicare and Medicaid Services (CMS), care for its beneficiaries with multiple chronic conditions accounts for 93% of Medicare fee-for-service expenditures. These patients often receive care from multiple physicians. A failure to coordinate care can often lead to patients not getting the care they need, receiving duplicative care, and being at an increased risk of suffering medical errors. On average, each year, one in seven Medicare patients admitted to a hospital has been subject to a harmful medical mistake in the course of their care. And nearly one in five Medicare patients discharged from the hospital is readmitted within 30 days – a readmission many patients could have avoided if their care outside of the hospital had been aggressive and better coordinated.

Improving coordination and communication among physicians and other providers and suppliers through Accountable Care Organizations will help improve the care Medicare beneficiaries receive, while also helping lower costs. (Health & Human Services Press Office, 2011)

Accountable care organizations

Under the proposed rule, an ACO refers to a group of providers and suppliers of services (e.g., hospitals, physicians, and others involved in patient care) that will work together to coordinate care for the patients they serve with Original Medicare (that is, those who are not in a Medicare Advantage private plan). The goal of an ACO is to deliver seamless, high quality care for Medicare beneficiaries. The ACO would be a patient-centered organization where the patient and providers are true partners in care decisions. (Health & Human Services Press Office, 2011)

How open advancement of EMR coding can help in ACO?

The key goal of ACO is to make medical providers and suppliers coordinate with each other to deliver seamless, high quality care for patients. In order to make this coordination happens among medical providers and suppliers, a necessary condition is that patients' EMR records should be interchangeable through all their EMR systems. Since every medical providers use different EMR systems, it could become a significant barrel during the ACO transition. Open Advancement of EMR Coding could help during this process. Since the very basic nature of EMR coding is to convert unstructured or semi-structured patient records which are composed of narrative text and other forms of records into structured records (codes). After patient records have been transformed into structured data, it is more possible for them to become interchangeable among different EMR systems.

Medical coding in China

Medical reform in China

In early 2009, the China Central Communist Party along with the China State Council announced a comprehensive healthcare reform initiative through a program titled “Opinions on Deepening Pharmaceutical and Healthcare System Reform.” In the wake of that initiative, China has taken numerous steps to further advance the development of its primary healthcare services system. By 2011, an additional 900 million Chinese will enjoy a basic level of care. Through its reform program, meanwhile, China will, by 2011, have extended its healthcare coverage to 90 percent of the country’s 1.3 billion people, thereby creating a primary healthcare service system that, for the first time ever, offers basic and affordable drugs to the masses. To ensure the success of its burgeoning healthcare system, the Chinese government has committed \$125B over the next three years. An estimated 46% of that will be allocated to medical insurance initiatives, 47% to healthcare provisions, and 7% to public health. (IMS Health, 2010)

According to the medical insurance initiatives, the insurance coverage of people in China will be estimated to reach over 90% in upcoming years. It is making a huge impact in every aspect of Chinese medical field, especially Chinese medical system. One part of Chinese medical insurance initiatives have indicated that on January 1st, 2012, Chinese government started to initialize the implementation of ICD-10 in medical coding and billing system from ICD-9-CM-3. Considering the drastically increasing complexity of ICD10 and the rapidly increasing demand of IT support due to the medical reform, Chinese healthcare system is facing similar challenges as U.S.

Similar to U.S medical coding and billing system, incorrect code assignment is usually caused by a misunderstanding of the diagnosis or procedure due to its complex nature. However, comparing to well-developed U.S medical coding and billing system with thousands of trained medical coders and the help of medical coding software, Chinese hospitals have to deal with much higher volume of insurance claiming cases with much less well-trained medical coders and fewer IT support.

After I did a survey about the current market of Chinese medical coding software, it seems that there are very few of Chinese companies have products in this category and the only type of the medical coding software available is the “lookup table” tool. I haven’t seen any company indicates the “automating translating” type of medical coding software as ProVation and UPMC TDC claimed they have. Considering the rapidly growing demands of Chinese medical coding support, I have a reason to consider this type of more advanced Chinese medical software may be a good opportunity for entrepreneurs in Chinese software market especially the ones with Natural Language Processing technology.

Natural language processing in Chinese

Despite its increasing importance, the uniqueness of Chinese renders the computer processing of the language distinctive and challenging. There are several organizations starting to provide Chinese NLP software as open source tools.

Stanford NLP research in Chinese

Stanford NLP works on a wide variety of research in Chinese Natural Language Processing, including word segmentation, part-of-speech tagging, syntactic and semantic parsing, machine translation, disfluency detection, prosody, and other areas.

Chinese word segmentation

It relies on a linear-chain conditional random field (CRF) model, which treats word segmentation as a binary decision task. It uses three categories of features: character identity n-grams, morphological and character reduplication features. As shown in the figure on the right, it also exploits lexicons and proper noun features to improve segmentation consistency, which is beneficial in tasks such as machine translation (MT) and information retrieval. *(This could be used for the Chinese Medical Coding Software)*

Parsing and grammatical relations

In addition to PCFG parsing, the Stanford Chinese parser can also output a set of Chinese grammatical relations that describes more semantically abstract relations between words. *(This could be used for the Chinese Medical Coding Software)*

Part-of-Speech tagging

The Stanford part-of-speech tagger takes word-segmented Chinese text as input and assigns a part of speech to each word (and other tokens), such as a noun or a verb. Its overall accuracy is 93.65% and the unknown word accuracy is 84.84%. *(This could be used for the Chinese Medical Coding Software)*

Tsinghua University NLP

The Natural Language Processing Group at the Department of Computer Science and Technology, Tsinghua University (THUNLP) is working on methodologies and algorithms for computer processing and understanding of human languages with emphasis on Chinese. It focuses on basic research in language computation as well as the application-oriented NLP technologies.

Its research covers a range of topics in natural language processing, including:

NLP based on huge-scale naturally annotated corpora

- Word segmentation using punctuations in huge-scale web articles
(This could be used for the Chinese Medical Coding Software)
- New word detection and related word retrieval from user logs of Chinese input method

- Chinese abbreviation extraction from anchor texts in webpages
(This could be used for the Chinese Medical Coding Software)
- New word detection from user logs of search engine

Social tagging and keyword extraction

- Tag disambiguation
- Tag suggestions using topic models
- Tag suggestions via Latent Reason Identification
- Exploring subsumption relations in social tags
- Keyword extraction by clustering to find exemplar terms
- Keyword extraction via topic decomposition

Multilingual analysis

- Fast and robust sentence alignment algorithm
- Bilingual terminology extraction system
(This could be used for the Chinese Medical Coding Software)
- Statistical method for Uyghur tokenization
- Uyghur morpheme analysis
- "Female Script" pinyin input method

Text classification

- Feature selection for Chinese text classification
(This could be used for the Chinese Medical Coding Software)
- Scalable term selection for text classification
- Efficient text classification using term projection
(This could be used for the Chinese Medical Coding Software)
- Transfer learning and self training classification
(This could be used for the Chinese Medical Coding Software)
- Text classification-based image classification for text

Companies involved in Chinese NLP

- Chinese Morphological Analyzer and Orthographic Converter
- Zi Corp: Makers of Chinese Input Technology
- Tegic Communications: Maker of Input Technology for Small Devices, specializing in Chinese
- Teragram Corporation: Makers of Language Software
- CJK Dictionary Publishing Society
- Inxight: Provides tools to view information.
- Microsoft's NLP Group
- Microsoft Research Beijing Lab
- BBN Technologies Speech and Language Processing Department
- SRA International does NLP-related work.
- Kent Ridge Digital Labs in Singapore.
- Systran Translation Software
- National Security Agency
- DARPA Information Technology Office: Sponsors a lot of the NLP research going on today.
- IBM's Research Lab in Beijing

Opportunities

Considering the variety of available Chinese Natural Language Processing technology that are developed by schools and organizations in China and the emerging demand for Chinese medical coding software due to the recent Chinese medical reform, I have reasons to believe that there are opportunities for entrepreneurs to bridge the gap by utilizing the right technology and solving the not-to-be-neglected problems. The future research may keep focusing on the Chinese medical coding market and Chinese Natural Language Processing technology.

Reference

- **Tribble, S. J.** (2012, 5 6). *Medical billing codes wreak havoc*. Retrieved 10 27, 2012, from The Columbus Dispatch: <http://www.dispatch.com/content/stories/business/2012/05/06/medical-billing-codes-wreak-havoc.html>
- **American Academy of Professional Coders.** (2012, January 1). 1. Retrieved September 15, 2012, from AAPC: <http://www.aapc.com/medical-coding.aspx>
- **American Medical Association.** (2012, January 1). *CPT - Current Procedural Terminology*. Retrieved September 15, 2012, from American Medical Association: <http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt.page>
- **American Medical Association.** (2008, October 1). *ICD-9-CM Official Coding Guidelines*. Retrieved September 25, 2012, from American Medical Association: http://www.ama-assn.org/resources/doc/cpt/icd9cm_coding_guidelines_08_09_full.pdf
- **Bureau of Labor Statistics.** (2012, August 7). 1. Retrieved September 15, 2012, from <http://www.bls.gov/>: <http://www.bls.gov/ooh/Healthcare/Medical-records-and-health-information-technicians.htm>
- **G&G Advanced Medical Consulting, Inc.** (2012, February 1). *Medical Billing Process*. Retrieved September 15, 2012, from Medical Billing Company: <http://www.medical-billing-company.com/medical-billing/medical-billing-process.html>
- **Health & Human Services Press Office.** (2011, March 31). *Affordable Care Act to improve quality of care for people with Medicare*. Retrieved October 5, 2012, from U.S. Department of Health & Human Services: <http://www.hhs.gov/news/press/2011pres/03/20110331a.html>
- **IMS Health.** (2010, 1 1). *Healthcare Reform in China*. Retrieved 12 12, 2012, from imshealth.com: http://imshealth.com/imshealth/Global/Content/Pharmerging/Document/China_Healthcare_Reform.pdf
- **International Health Terminology Standards Development Organization.** (2012, July 1). *SNOMED CT User Guide*. Retrieved September 30, 2012, from Veterinary Terminology Services Laboratory: http://vtsl.vetmed.vt.edu/Education/Documentation/SNOMED/doc_UserGuide_Current-en-US_INT_20110731.pdf
- **PrecisionCode.** (2012, 1 1). *PrecisionCode*. Retrieved 10 21, 2012, from MDsolutions: <http://www.mdsolutions.com/ecode.html>
- **Rouse, M.** (2010, February 1). *Current Procedural Terminology*. Retrieved September 29, 2012, from Search Health IT: 13957305689
- **Spinello, S.** (2009, January 1). *CPT for Medical Procedure*. Retrieved September 25, 2012, from eHow: http://www.ehow.com/about_5318785_cpt-medical-procedures.html

- **Torrey, T.** (2012, July 17). *What are ICD-9 or ICD-10 codes?* Retrieved September 15, 2012 , from Patient Empowerment:
<http://patients.about.com/od/medicalcodes/a/icdcodes.htm>
- **Tu, J. I.** (2012, 9 21). *Microsoft, GE uniting to create better health-care data systems.* Retrieved 10 26, 2012, from Seattle Times.
- **White Plume.** (2012, 1 1). *AccelaSMART - An Intelligent Management and Workflow Engine.* Retrieved 10 21, 2012, from White Plume:
<http://www.whiteplume.com/products/accelasmart-an-intelligent-management-and-workflow-engine>