Using Natural Language Processing To Analyze Electronic Health Records

Philip Poon PhD
Data Scientist

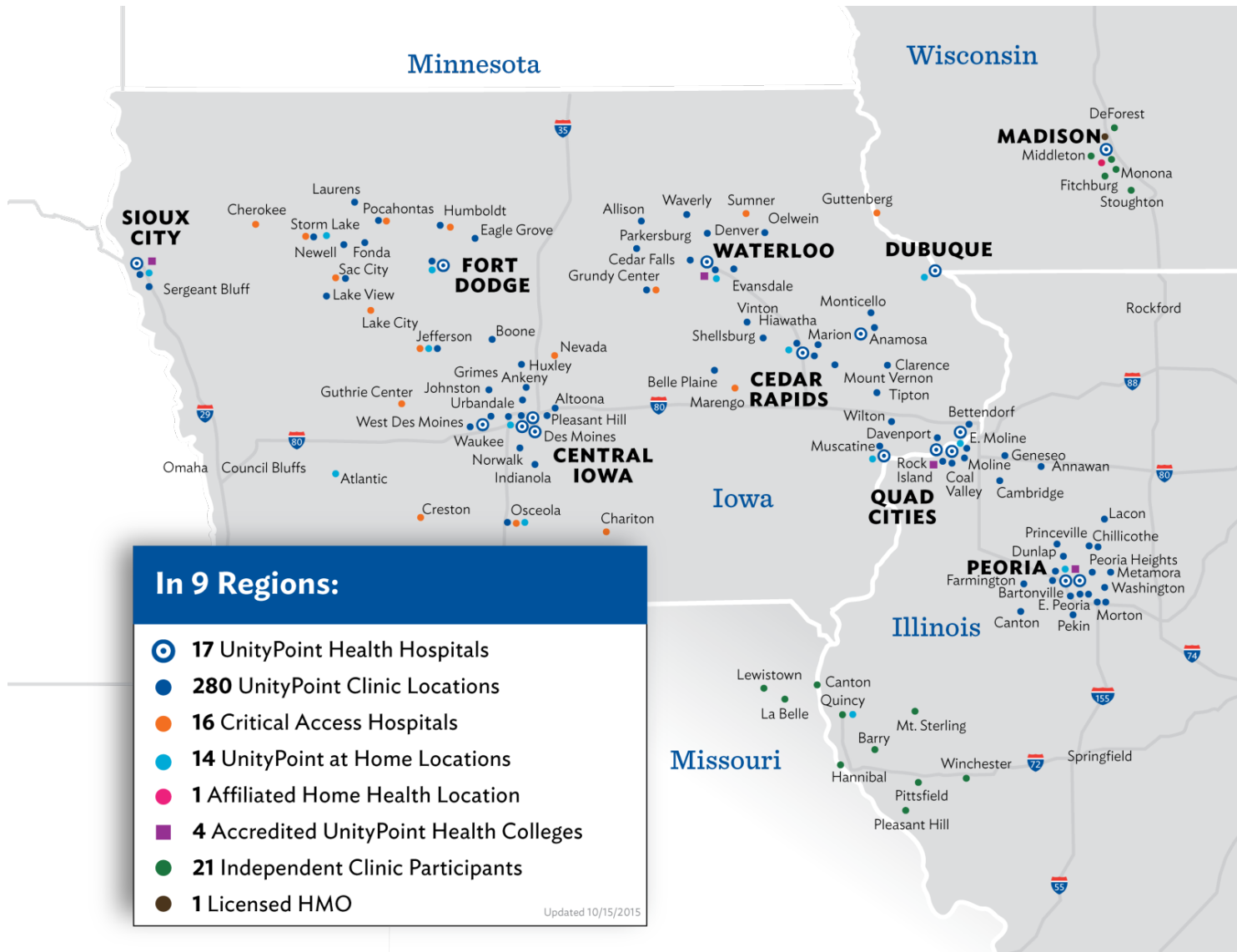UnityPoint Health

Agenda

- Overview of UnityPoint Health and Electronic Health Record

- Natural Language Processing development

- Current Use Cases and Results

UnityPoint Health

Overview

- Electronic Health Record (EHR) - Epic
  - Software used to document care provided
  - Accessible by all providers in UnityPoint Health (UPH)
  - Contains structured (lab tests, medications, height, etc.) and unstructured data (physician notes, progress reports, etc.)
  - 21.5M Free-Text Notes from 2015 (42GB)

- Natural Language Processing (NLP) is applied to turn free text into usable features for modeling and analysis

UnityPoint Health

UnityPoint Health

In 9 Regions:

- ⊙ **17** UnityPoint Health Hospitals
- ● **280** UnityPoint Clinic Locations
- ● **16** Critical Access Hospitals
- ● **14** UnityPoint at Home Locations
- ● **1** Affiliated Home Health Location
- ■ **4** Accredited UnityPoint Health Colleges
- ● **21** Independent Clinic Participants
- ● **1** Licensed HMO

Updated 10/15/2015

UnityPoint Health

- **3788 Beds**

- **155,334 Inpatient visits**
  - **Hospital or Rehabilitation Stays**

- **5.4M Outpatient visits**
  - **Clinic visits**

- **900+ Physicians**

- **30000+ Employees**

UnityPoint Health

IBM Watson

- Corpus: 200 million pages of structured and unstructured content (for Jeopardy)
- Including full copy of Wikipedia
- Not connected to the internet
- Has been commercialized for use in healthcare
- Build medical-specific terms into dictionaries.
- Develop character rules to recognize patterns of characters that represent specific types of information, (e.g. telephone numbers or vital measures).
- Develop parsing rules to identify patterns of text that represent particular concepts, such as person names.
- Unstructured Information Management Architecture (UIMA) pipeline directly to Watson NLP server and automatically create fields and facets that are mapped to the annotations.
- Analyze documents with a Watson Content Analytics pipeline and view the resulting annotations in ICA Studio.

UnityPoint Health

Development Cycle
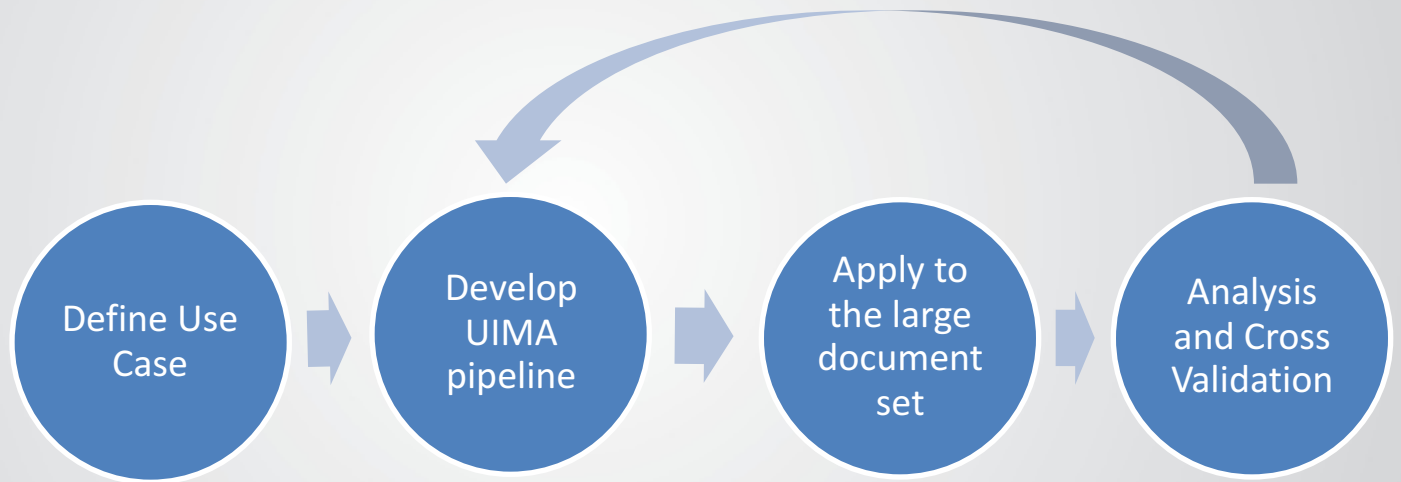
An Example

"...JANE is taking her medications as prescribed.Active Problems Allergic rhinitis (477.9) Benign essential hypertension (401.1); JNC 8 goal < 150/90 Degenerative disc disease (722.6); L4-L5 and L5 S1 Depression (311) Dyslexia (784.61) Fusion of spine of lumbar region (756.15) GERD (gastroesophageal reflux disease) (530.81) Glucose intolerance (impaired glucose tolerance) (790.22) Hip osteoarthritis (715.95) Hyperlipidemia (272.4) Hypothyroidism (244.9) Ventral hernia (553.20).PMH History of anemia (V12.3) History of migraine headaches (V12.49).PSH Arthrodesis Lumbar Colonoscopy (Fiberoptic); 2006 Dilation And Curettage Hysterectomy Tonsillectomy Wrist Surgery.Family Hx Benign essential hypertension: Mother,Father,Sibling Family history of alcoholism: Sister,Brother (V17.0) Family history of cataracts: Sister (V19.19) Family history of diabetes mellitus: Mother,Father,Sibling (V18.0) Family history of glaucoma: Sister (V19.11) Family history of malignant neoplasm of urinary bladder: Aunt (V16.52).Personal Hx Lives with parents Never a smoker..."

UnityPoint Health

An Example

- **Active problem: Benign essential hypertension (401.1), impaired glucose tolerance (709.22)...**
- **PMH: magraine (V12.49)...**
- **PSH: Arthrodesis Lumbar Colonoscopy (?)**
- **Family Hx: Benign essential hypertension: Mother,Father,Sibling; diabetes mellitus: Mother,Father,Sibling (V18.0).**
- **Personal history: Lives with parents. Never a smoker.**

UnityPoint Health

Another Example

"Have not been able to contact patient by phone 3 times. Per protocol, patient is dropped from COPD education program."

Does this patient has COPD?

Yes!

UnityPoint Health

Challenges

- **Coding Change from ICD9 to ICD10 on Oct 2015.**

  1000 unique entries to 24000.

- **Messy data (symbols, formatting, misspelled terms)**

  e.g. •, diabetes melitus...

UnityPoint Health

Use Case: Chronic Conditions

- **Does our documentation accurately reflect a patient's complexity?**
  - **Affects how much we're reimbursed for providing care**

- **Example: Diabetes**
  - **Identified about 10% of 330,000 Medicare patients were not properly documented as diabetics**
  - **Proper documentation would have resulted with big revenue saving.**

UnityPoint Health

Use Cases: Social Determinants

- **Living arrangements, education, employment, income…**

- **Can be used in building predictive models and providing actionable information for care givers**

UnityPoint Health

Thanks!

Philip Poon

philip.poon@unitypoint.org