What's in a narrative note? Data from i2b2 shared tasks 2007-2012
Anna Rumshisky, Özlem Uzuner
MIT Computer Science and Artificial Intelligence Laboratory

WHAT'S IN A NOTE?
- Information recorded by clinicians in narrative notes is not readily available for search and retrieval.
- Automatic indexing and extraction of such information requires training the software on text annotated by humans.

SHARED TASKS
- A series of annotation efforts have been carried out by i2b2 from 2007 to 2012, covering a range of information contained in the text of clinical notes.
- A system development competition (“challenge” / “shared task”) have been run in conjunction with these efforts in order to encourage the development of automated information extraction software.
- Creating a “layered” linguistic annotation, where each record the data set is annotated with several layers of linguistic and clinical information.

DATA SETS
Data sets created and the corresponding system development tasks:
- De-identification of clinical records (2007)
- Smoking status detection at the patient level (2007)
- Record-level identification of obesity and its 15 co-morbidities (2008)
- Extraction of medications and the associated attributes, including dosage, frequency, etc. (2009)
- Extraction of medical problems, tests, and treatments from discharge summaries (2010)
- Identifying the assertion status of medical problems (presence, absence, attribution, etc.) (2010)
- Extracting relationships between medical problems, treatments, and tests (administered for, etc.) (2010)
- Co-reference relations between medical problems, treatments, tests, and people, including resolution of pronouns (2011)
- Sentiment detection in suicide notes (2011)
- Temporal relations between events and event anchoring to dates, times, durations and frequencies (2012)

1. DE-IDENTIFICATION
- ~900 discharge summaries from Partners Healthcare
- Private health information (PHI) as defined by HIPAA
- 8 categories of PHIs marked:
  - Patients, doctors, hospitals, ids (including medical record, device, license, and other ids), dates, locations, phone numbers, and ages over 90 yo.
- 7 teams developed systems for the task
- Best-performing systems scored 98% f-measure on all PHI categories

2. SMOKING STATUS
- ~500 de-identified discharge summaries from Partners Healthcare
- Records classified by pulmonologists into five categories:
  - Past smoker, current smoker, smoker (unclear if past or current), non-smoker (never smoked), and unknown (smoking status unknown)
- 11 teams developed systems for the task
- Microaveraged f-measure above 84%

3. OBESITY AND ITS CO-MORBIDITIES
- ~1250 discharge summaries from Partners RPDR
- Records classified by experts w.r.t to obesity and its 15 co-morbidities:
  - Asthma, atherosclerotic cardiovascular disease (CAD), congestive heart failure (CHF), depression, diabetes mellitus (DM), gallstones / cholecystectomy, gastroesophageal reflux disease (GERD), gout, hypercholesterolemia, hypertension (HTN), hypertriglyceridemia, obstructive sleep apnea (OSA), osteoarthritis (OA), peripheral vascular disease (PVD), and venous insufficiency
- 4 categories:
  - Present (Y): the patient has the disease; Absent (N): the patient does not have the disease; Questionable (Q): the patient may have the disease; Unmentioned (U): the disease is not mentioned in the record
- 2 judgments: textual (based on explicit language in the summary) and intuitive (requiring expert inference)
- 30 teams developed systems

4. MEDICATIONS AND THEIR ATTRIBUTES
- ~1250 discharge summaries from Partners Healthcare
- Medications and their attributes:
  - Medications (brand names, generics, collective names of prescription substances); Doses; Modes (routes); Frequencies; Durations; Reasons; whether they appeared in a list or in a narrative segment
- Strings and offsets matching each attribute
- 20 teams developed systems
- For top 10 systems, 75% f-measure for medications, 53% for durations and 46% for reasons

5. MEDICAL PROBLEMS, TREATMENTS, TESTS, AND RELATIONS BETWEEN THEM
- ~900 discharge summaries and progress notes from Partners Healthcare, Beth-Israel Deaconess Medical Center, and U. Pittsburgh Medical Center
- Three information extraction tasks:
  - Extracting clinically relevant concepts, including (1) medical problems, (2) treatments, (3) tests
  - Identifying assertion status of the mentioned medical problems as: present, absent, possible, conditional, hypothetical, and attributed to someone else
  - Identifying relations between clinical concepts: treatments being administered for, treatment improving, worsening, or causing a medical problem; tests revealing or being conducted to investigate a medical problem, medical problems indicating other medical problems
  - e.g. He is status post radiation [treatment] for non Hodgkin’s lymphoma [problem] → treatment administered for problem
- Concept extraction: 22 teams, with 74% mean exact f-measure. Assertion status identification: 21 teams, with 86% mean f-measure. Relation extraction: 16 teams, with 60% mean f-measure

8. CO-REFERENCE RELATIONS
- ~800 discharge summaries and progress notes from Partners, BIDMC and U Pittsburg MC
- Co-reference resolution: Identifying mentions referring to the same entity
- Resolution of coreference between mentions of medical problems, treatments, tests, and people, including pronouns
  - e.g. tap | the procedure | it [treatment]
  - BB | beta-blocker | Coreg [treatment]
- 16 teams submitted systems, with average f-measure in 70-85% for diff. tasks and datasets

9. TEMPORAL RELATIONS
- ~300 discharge summaries from Partners RPDR
- Clinically relevant events (including admission, discharge, transfers between clinical departments, tests, procedures, symptoms, etc.): temporal expressions
  - including times, dates, durations, and frequencies,
  - and relations between them (before, after, simultaneous, etc.)

PRACTICAL APPLICATIONS: What questions can we now answer?
- Given the layered linguistic annotations on which information extraction systems have been trained, what clinically useful retrieval tasks can we now facilitate?
- Reasons for admission by diagnosis? Temporal and causal relations between symptoms and problems? Medication timelining? Relations between specific tests and treatments, medications and specific procedures?