

HARVEST, a Holistic Patient Record Summarizer at the Point of Care

Noémie Elhadad, PhD¹, Sharon Lipsky Gorman, MA¹, Jamie S. Hirsch, MD, MS¹,
Connie Liu, MS¹, David K. Vawdrey, PhD¹, Marc Sturm²

¹Columbia University, New York, NY; ²NewYork-Presbyterian Hospital, New York, NY

Introduction

In many care settings, clinicians are faced with an overwhelming amount of complex information about their patients, with little time for chart review [1]. Failure to digest patient data may result in errors in diagnosis and delayed care [2,3]. The need for better health information management and visualization tools has long been recognized [4,5], yet current electronic health records (EHRs) do not yet provide the cognitive support necessary for effectively and efficiently reviewing patient data. EHR displays are plagued by low information content, do not honor established interface design principles, and cannot be readily customized without imposing a considerable burden on clinicians and information technology professionals [6,7]. Important research has been undertaken on visualization of patient histories [8–10] and domain-specific summarizers [11], motivating our hypothesis that a holistic patient-record summarizer can impact care in a beneficial fashion. In this abstract, we describe the design, implementation, and deployment of HARVEST, a longitudinal patient record summarization system.

Methods

HARVEST is an interactive, problem-oriented patient record summarization system (see Figure below) [12]. The summarizer differs from previous work in three critical ways: (i) it extracts content from the patient notes, where key clinical information resides; (ii) it aggregates and presents information from multiple care settings, including inpatient, ambulatory, and emergency department encounters; and (iii) it is integrated into two commercial EHR systems, and is available for all patients in our institution, not just a curated dataset or for specific patient cohorts.

The natural language processing (NLP) of clinical notes is carried out through a named-entity recognition system that indexes concept mentions. Because HARVEST aggregates problems, NLP was constrained to extracting concepts from the UMLS semantic group “Disorder” only, restricted to the SNOMED-CT Core Problem List. Concept salience weights are computed dynamically to reflect both the frequency of the concept in the patient notes in a given time slice of the record, and the prevalence of the concept across all patients in the institution. To enable parsing and salience computation at scale, we created a distributed computing infrastructure (using Apache Hadoop) and implemented a map-reduce version of our NLP system to parse the notes from a variety of HL7 interface feeds.

Results

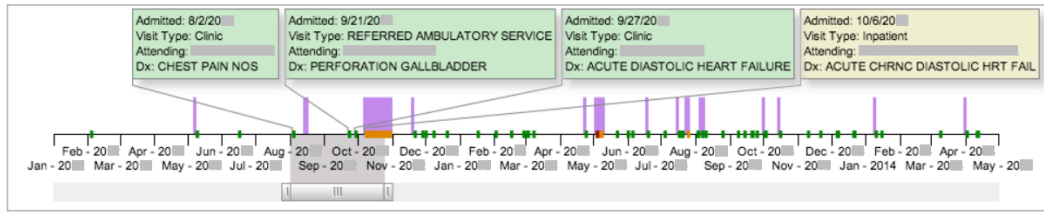
The Hadoop infrastructure enabled us to accommodate the large volume of clinical documents generated in our institution (650,000 notes per month): a small four-node cluster processed 20,000 notes/second compared to 500 notes/second in a non-distributed computing environment. The infrastructure also permitted us to experiment with different parsing and salience computation strategies through short development cycles.

HARVEST was deployed within NewYork-Presbyterian’s iNYP clinical information review system as a beta release in September 2013. As a first phase of deployment and to study the impact of HARVEST on clinical care, access was limited to internal medicine residents and residents and attending physicians in the emergency department (ED). Three primary use cases for the summarizer have emerged thus far: (i) in the ED, as a way to capture the essential knowledge about a patient’s history in an efficient fashion, including previous emergency visits and patterns of visits; (ii) in internal medicine, as a support tool for performing an in-depth patient chart review before admitting a patient to the hospital; and (iii) in the primary care clinic, for chief residents as an education tool during precepting hours with residents.

Discussion

HARVEST addresses an unmet need for clinicians at the point of care, facilitating effective and efficient review of essential patient information. The deployment of HARVEST in our institution allows us to study patient record summarization as an informatics intervention in a real-world setting. It also provides an opportunity to learn how clinicians use the summarizer, enabling informed interface and content iteration and optimization to improve patient care. Future work includes customizing the content selected by HARVEST for different types of clinicians.

Timeline: 8/1/20 to 10/24/20



stable angina pulmonary hypertension ESRD **dyspnea** influenza abdominal pain DM CAD
 edema volume overload obese OSA chest pain lymphadenopathy morbid obesity pruritis weight gain hypertension DM2 LVH
 leg cramps chest discomfort vitamin D deficiency CKD hyponatremia agitation fistula nausea facial swelling hypoglycemia ischemia
 CHF Dyslipidemia abdominal mass scar hyperphosphatemia anasarca angina hypoventilation ...

Notes about dyspnea 8/1/20 - 10/24/20		Cardiology Consult Free Text Note
Cardiology Consult Follow-up Free Text Note	10/15/20 1:32 PM	Cardiology Consult Requested by: Dr. [REDACTED] Reason: Fluid overload HPI: 57 yo woman with a pmhx significant for morbid obesity, HTN, HLD, DM2, CKD (stage V) not on RRT and making urine, CAD s/p mLAD DES in 7/20, and pulmonary HTN (based on RHC on 7/20) who presents with signs and symptoms of fluid overload. Cardiology is being asked to consult for further management. In regards to the patient's functional status, the patient lives a sedentary lifestyle and is now on disability. Over the course of the past month, she has had increasing fluid accumulation with a weight gain of over 25 kg, with worsening LE edema and facial puffiness. Prior to 1 month ago, her ET was 2 blocks, but has now decreased to 15 feet limited by SOB and occasionally with CP. Furthermore, she has a 6 pillow orthopnea that has been stable for 4 years but has had worsened PND this past month. The patient also reports 3 months of intermittent chest pain. She describes the pain as sharp, retrosternal, and located in the center of the chest, lasting 5 minutes with 1-2 episodes per week. These episodes occur at rest, and improved by sitting up and taking an aspirin. PMHx: 1. Morbid obesity 2. HTN 3. HLD 4. DM2
Milstein Hospitalist Resident/PA Follow-up Free Text Note	10/15/20 7:00 AM	
Medicine Follow-Up Free Text Note	10/14/20 4:06 AM	
Nephrology Consult Free Text Note	10/13/20 2:52 PM	
Milstein Hospitalist Attending Follow-up Free Text Note	10/13/20 11:27 AM	
Cardiology Consult Follow-up Free Text Note	10/12/20 11:40 AM	
Milstein Hospitalist Resident/PA Follow-up Free Text Note	10/12/20 7:02 AM	
Milstein Hospitalist Resident/PA Follow-up Free Text Note	10/11/20 12:43 PM	
Cardiology Consult Free Text Note	10/10/20 10:14 AM	
Medicine Follow-Up Free Text Note	10/10/20 10:10 AM	
Case Manager Plan of Care	10/10/20 5:31 AM	
Milstein Hospitalist Resident/PA Follow-up Free Text Note	10/09/20 7:58 AM	
Milstein Hospitalist Resident/PA Follow-up Free Text Note	10/08/20 7:21 AM	
Nursing Adult Admission History	10/07/20 2:24 AM	
Medicine Admission Free Text Note	10/06/20 11:30 PM	
ED Resident/NP/Attending Note (Milstein)	10/06/20 3:04 PM	

panel, with all mentions of dyspnea (and synonyms) highlighted. On the timeline, documentation of dyspnea is highlighted by purple bars, indicating that dyspnea was a particularly salient issue at that time, as well as 6 months later.

Figure 1. De-identified HARVEST screenshot for a sample patient, part of iNYP. For the selected time range (3 months), stable angina, pulmonary hypertension, end-stage renal disease, and dyspnea are the most prominently documented problems. HARVEST also identified diabetes mellitus, hypertension, and dyslipidemia as important problems. The problem “dyspnea” was selected. The Notes panel lists all notes in the selected time range that mention this problem. A cardiology consult note is selected and displayed in the lower right

References

- Reichert D, Kaufman D, Bloxham B, et al. Cognitive analysis of the summarization of longitudinal patient records. In: *Proceedings of the Annual American Medical Informatics Association Fall Symposium (AMIA)*. 2010. 667–71.
- Christensen T, Grimsmo A. Instant availability of patient records, but diminished availability of patient information: a multi-method study of GP’s use of electronic patient records. *BMC Med Inform Decis Mak* 2008;8:12.
- Singh H, Spitzmueller C, Petersen NJ, et al. Information overload and missed test results in electronic health record-based settings. *JAMA Intern Med* 2013;173:702–4.
- Powsner SM, Tuft ER. Graphical summary of patient status. *Lancet* 1994;344:386–9.
- Payne TH. Computer decision support systems. *Chest* 2000;118:47S–52S.
- Blumenthal D, Tavenner M. The ‘meaningful use’ regulation for electronic health records. *N Engl J Med* 2010;363:501–4.
- Laxmisian A, McCoy AB, Wright A, et al. Clinical summarization capabilities of commercially-available and internally-developed electronic health records. *Appl Clin Inform* 2012;3:80–93.
- Plaisant C, Mushlin R, Snyder A, et al. LifeLines: using visualization to enhance navigation and analysis of patient records. In: *Proceedings of the Annual American Medical Informatics Association Fall Symposium (AMIA)*. 1998. 76–80.
- Shahar Y, Boaz D, Tahan G, et al. Interactive Visualization and Exploration of Time-oriented Clinical Data Using a Distributed Temporal-Abstraction Architecture. In: *Proceedings of the Annual American Medical Informatics Association Fall Symposium (AMIA)*. 2003. 1004.
- Tao C, Wongsuphasawat K, Clark K, et al. Towards event sequence representation, reasoning and visualization for EHR data. ACM Press 2012. 801.
- Wilcox A, Jones SS, Dorr DA, et al. Use and Impact of a Computer-Generated Patient Summary Worksheet for Primary Care. In: *Proceedings of the Annual American Medical Informatics Association Fall Symposium (AMIA)*. 2005. 824–8.
- Weed LL. Medical records that guide and teach. *N Engl J Med* 1968;278:652–657 concl.