

NLP-Based Radiology Report Simplification and Summarization System

Spencer Beck, MD

Pratheek Bobba, BS

Sophie Chheang, MD, MBA

Ali Mozayan, MD

Dragomir Radev, PhD

Anne Sailer, MD

Keen You, BS

Background

- Natural language processing (NLP): artificial intelligence applied to the computer interpretation of human language
- Increasingly robust computer processing resources, radiological data, and machine learning algorithms advance the potential of NLP
- Radiology reports employ heavy use of jargon
- Patients on average have an eighth-grade reading level
- 40% of radiology report recommendations are not followed through
- 21st Century Cures Act: No delay in electronic access to clinical information, including radiology reports, to patients



Why Radiology Reports

- Means of inter-physician and physician-to-patient communication
- Unique lexicon derived from the imaging modality being utilized and the body part being studied
 - Facilitates use of pre-trained models
- Structured reporting including subsections for findings and impression
- Established databases of radiology reports to train and test a summarization and simplification model
- Reports take less system resources than images to store and access
- Simplified report would assist clinicians in explaining imaging findings to patients
 - This in turn would increase patient engagement and likelihood of follow-through on clinician recommendations



Our Team

- Recognized experts in radiology and natural language processing, with high-impact scientific publications
- Unique collaboration between Yale School of Medicine and Yale Computer Science leading to development of an NLP-based radiology report simplification system



Sophie Chheang, MD, MBA
Assistant Professor of Radiology
Assistant Director of Informatics

Department of Radiology and
Biomedical Imaging



Drago Radev, Ph.D.
A. Bartlett Giamatti Professor of
Computer Science
LILY (Language, Information, and
Learning at Yale) lab

Department of Computer Science



Ali Mozayan, MD
Radiology Resident



Keen You
Computer Science
PhD Candidate



Anne Sailer, MD
Radiology Resident



Spencer Beck, MD
Radiology Resident



Pratheek Bobba
Medical Student

RSNA Journals CME Contact Us Subscribe E-mail Alerts

RadioGraphics

Latest Articles | Current Issue | All Issues | Collections | Authors/Reviewers | Core

Informatics

Practical Guide to Natural Language Processing for Radiology

Ali Mozayan, Alexander R. Fabbri, Michelle Maneevese, Irena Tocino, Sophie Chheang

Author Affiliations

Published Online: Sep 1 2021 | <https://doi.org/10.1148/rg.2021200113>

JAIR Journal of Artificial Intelligence Research

CURRENT ARCHIVES ABOUT SPECIAL TRACKS

Home / Archives / Vol. 22 / Articles

LexRank: Graph-based Lexical Centrality as Saliency in Text Summarization

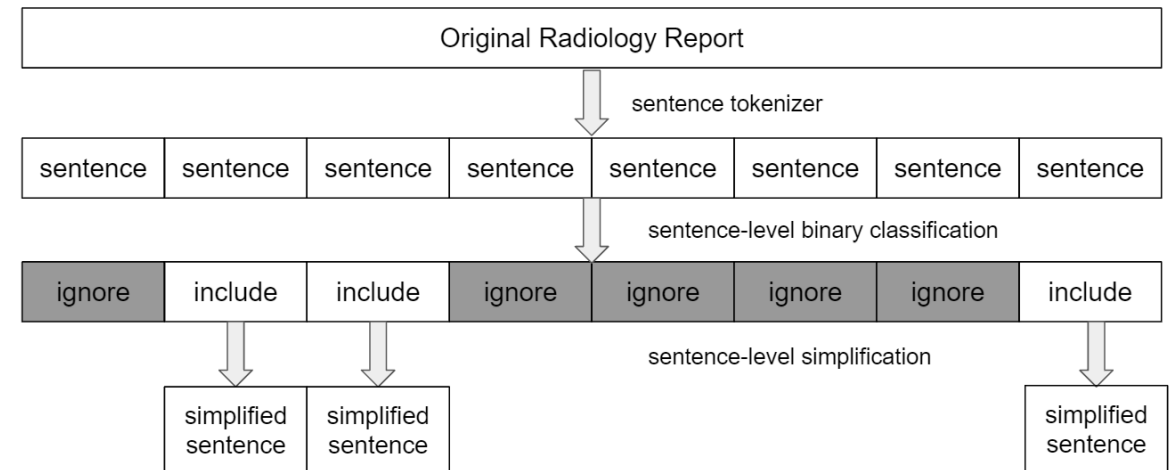
G. Erkan, D. R. Radev



Under the Hood

- Framework: Bidirectional Encoder Representations from Transformers (BERT)
 - Open-source, unsupervised framework published by Google AI in 2018
 - Transformer-based machine learning technique for NLP
- Omits normal findings
- Omits findings with little medical significance
 - Anatomic variants (two vessel aortic arch)
 - Benign incidental findings (simple renal/liver cysts)
- Simplifies medical jargon and spells out abbreviations
 - Pleural effusion → fluid around the lungs
 - Metastatic disease → cancer spread

Task Overview: Simplification Pipeline



Actual Model-Generated Simplified Radiology Report

Lungs/Airways/Pleura : Granulomas again seen. No new or enlarging pulmonary nodules. Interval resolution of pulmonary edema and bilateral pleural effusions. Central airways are patent.

Mediastinum/Lymph nodes: No thoracic adenopathy.

Heart and Vessels: The heart is stable in size. There is mild coronary artery calcification. No pericardial effusion. Right IJ port terminates in the mid to lower SVC.

Upper Abdomen: The abdominal CT portion of this study will be dictated separately by the body service.

Osseous structures and Soft Tissues: No aggressive bone lesions. Degenerative changes are seen in the visualized spine. Again seen are old left rib fractures.

Fluid in the lungs and fluid around the lungs are no longer seen.

There is mild hardening of the heart's blood vessels.



Age-related changes are seen in the visualized spine.

Again seen are old left rib fractures.

Market Analysis and Competition

Adoption Rate	Low	Medium	High
# of U.S. hospitals	6,000		
Client base	3%	5%	7%
Annual license fee	\$100 K	\$150 K	\$200 K
Annual Revenue	\$18 M	\$45 M	\$84 M

To our knowledge, no dedicated radiology report simplification or summarization tool exists. A start-up, Scanslated, offers hyperlinked radiology reports leading to a visual dictionary of medical terms, but does not actually simplify the reports directly.

Some non-medical websites, such as Simple English Wikipedia and Newsela, offer versions of the original document in different reading levels.

Patients currently use online search engines, various patient-centered health websites, and their personal radiologist network to get help understanding their radiology reports.



How Blavatnik Funding Will Get Us There

- Employ additional medical staff to manually summarize reports to train the system
 - Current model is derived from only 189 simplified radiology reports
 - We believe accuracy will be greatly improved with more training data
- Expand into other imaging modalities and body parts as current focus is on Chest CT reports
- Conduct A-B testing
 - Referring provider satisfaction
 - Patient satisfaction
 - Patient adherence to follow-up recommendations and projected cost savings to the healthcare system
- Explore feasibility of separate billing code for simplified report
 - Increase billing potential and adoption rate

We believe that routine simplification and summarization of radiology reports will result in

1. Better clinical outcomes
2. More efficient delivery of healthcare
3. Reduced downstream costs

