NTCIR-17 kickoff event
Sept 28, 2022

NTCIR-17 Medical Natural Language Processing for Social media and Clinical texts (MedNLP-SC)

Eiji Aramaki, Ph.D. (NAIST, Japan) Co-chair (general)
Shoko Wakamiya, Ph.D. (NAIST, Japan) Co-chair (general)
Shuntaro Yada, Ph.D. (NAIST, Japan) Co-chair (Subtask SM)
Yuta Nakamura, M.D. (The University of Tokyo, Japan) Co-chair (Subtasks RR)

MedNLP series

Name	Text	Task	Languages
NTCIR-10 MedNLP pilot	Dummy health records (by hand)	NER	JA
NTCIR-11 MedNLP-2	Dummy health records (from examinations)	NER, NEN	JA
NTCIR-12 MedNLPDoc	Dummy health records (from the textbook)	ICD coding	JA
NTCIR-13 MedWeb	Dummy social media texts (created by crowdsourcing)	Phenotyping	JA EN ZH
NTCIR-16 Real-MedNLP	Real health records: Case reports and Radiology reports	NER, ADE, CI	JA EN
NTCIR-17 MedNLP-SC Social Media (SM)	**Dummy social media texts (SM)**	**ADE**	**JA EN FR DE**
NTCIR-17 MedNLP-SC Radiology Report (RR)	**Real health records: Radiology reports (RR)**	**NER, TNM**	**JA**

Named Entity Recognition (NER)
Named Entity Normalization (NEN)
Adverse Drug Event detection (ADE)
Case Identification (CI) (MedNLP original task)
TNM cancer staging (TNM) (MedNLP original task)

MedNLP-SC subtasks

Information extraction from medical reports written by physicians or social media data

- **Subtask for Social Media (SM)**
  - Adverse drug event detection (ADE)
  - Text: Social media data (n=20,000 artificial tweets)
  - Language: Japanese  , English  , French  , and German 
- **Subtasks for Radiology Report (RR)**
  - (a) Named Entity Recognition (NER)
  - (b) TNM staging (TNM)
  - Text: Radiology reports (n=135 reports or more)
  - Language: Japanese 

Subtask for Social Media (SM)

Social media data (tweets)

- Create a Japanese corpus consisting of 20,000 artificial tweets
  - Out of a maximum of 68,000 artificial tweets (= 68 drugs x 1000 tweets) generated using T5
  - Each tweet is manually checked and annotated symptom label(s)
- Translate the corpus into the other three languages by machine translation with manual check
  - All language shares the same symptom label(s)

Samples of symptom labels

Label	Japanese	English	French	German
PAIN	疼痛	pain	douleur	Schmerzen
CONSTIPATION	便秘	constipation	constipation	Verstopfung
DIARRHEA	下痢	diarrhea	diarrhée	Diarrhöe
NAUSEA	吐き気	nausea	nausées	Übelkeit
VOMITING	嘔吐	vomiting	vomissements	vomissements

Subtask for Social Media (SM)

Artificial tweet samples of "**gamma globulin**" (1/2)

- Example 1

  検索もせずただひたすら我慢してるだけマシなんだが…。**γグロブリン**大量投与で一発逆転できるのかしら、それとも**ガンマグロブリン**大量投与か……ってなると何回やっても思い出せない

  (I'm better off not searching for it and just holding on... I wonder if a large dose of **gamma globulin** would turn it around, or a large dose of **gamma globulin**. ..... I can't remember how many times I've tried.) translated by DeepL

- Example 2: No drug name

  持病の関係で昨日病院行ったら血液検査の結果が出て「リンパ球減少症」と診断されました¥nでも今日からまた薬が増えます〜!頑張れ私!!'

  (I went to the hospital yesterday because of my chronic illness and the results of my blood test came back and I was diagnosed as "lymphopenia." Hang in there me!)

  translated by DeepL

Subtask for Social Media (SM)

Artificial tweet samples of "**gamma globulin**" (2/2)

- Example 3: No ADE

  けど、このところ連日の高熱が続きすぎて辛いです…今夜も頑張って寝ます!(*^^*)今日は病院に行って血液検査と**γグロブリン**注射をしてきました。これで少しでも良くなりますように!!お騒がせしましたm(=_=;")早く元気になって欲しいものです

  (But I'm having a hard time because I've had too many high fevers in a row recently... I will try my best to sleep tonight! (*^^*) I went to the hospital today for a blood test and a **gamma globulin** shot. I hope this will make me feel a little better! Sorry for the trouble m(=_=;")I hope she gets well soon!) translated by DeepL

- Example 4: Just mentioned

  <user_name> さん、こんにちは。私も昨日2回目打ってきましたよ!¥n接種後から頭痛と熱っぽさが出ていたので**γグロブリン**製剤を飲むように指示されたそうです。'

  (Hello, <user_name>. I got my second shot yesterday too! ¥They told me to take a **gamma globulin** preparation because I had a headache and feverishness since after the on vaccination.') translated by DeepL

Subtask for Social Media (SM)

Adverse Drug Event detection (SM-ADE)

- Multi-labeling task to identify a set of symptoms caused by a drug (section)

Given an input sentence in the section of FU5:
"After FU5 starts, I am suffering from dry cough and many mouth ulcers."
The expected output labels:
"dry cough" and *"mouth ulcers"*

- Similar to hashtag recommendation because symptom label(s) can be represented as hashtags

  e.g., *"After FU5 starts, I am suffering from dry cough and many mouth ulcers. #dry_cough #mouth_ulcers"*

FAQ for Social Media Subtask

(a) The performance of T5 to create tweet-like short messages

Answer: To ensure that the artificial tweets are sufficiently similar to real ones, we will conduct the Turing Test, which annotators determine whether a tweet is real or artificial

(b) The quality of the artificial tweets

Answer: To keep the quality of our dataset, artificial tweets will be checked by annotators and modified/removed if not appropriate

(c) The appropriateness of applying machine translation

Answer: Due to the cost (both time and money), we will employ machine translation

(d) Evaluation methods

Answer: We plan to employ F-measure

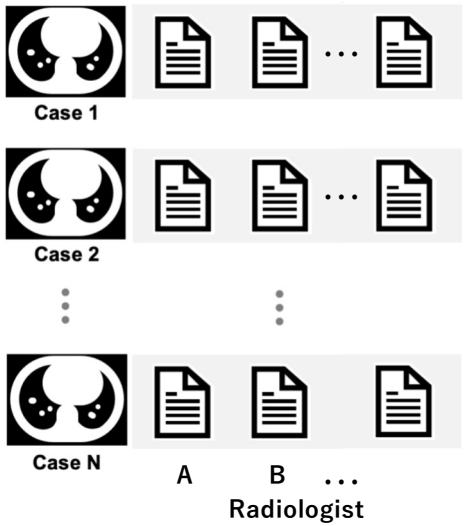
Subtasks for Radiology Report (RR)

image: https://radiopaedia.org/cases/cavitating-lung-cancer-1
Report: J-MedStd RR

MedTxt-RR Corpus

A set of radiology reports of the same cancer cases diagnosed by different radiologists independently

- The latest version contains 135 texts for 15 cases and 9 radiologists, some of which are already available at https://sociocom.naist.jp/medtxt/rr/
- We are planning to make the dataset larger by recruiting additional cancer cases and radiologists



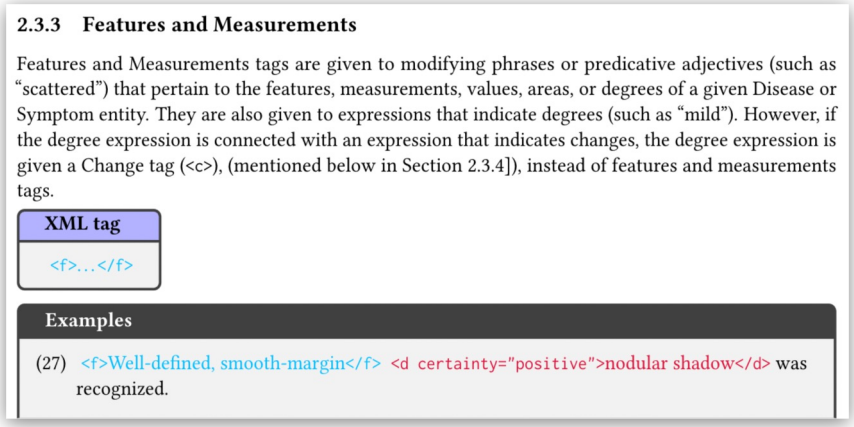
Radiologist A's report	Radiologist B's report
両肺には軽度の気腫性変化、左肺上葉外側にbullaを認めます。	
左肺下葉には内部に空洞を伴う長径7.8cm 大の不整形腫瘤を認め、	左肺下葉に長径78mm大の腫瘤があり，既知の肺癌と考えます。
同一肺葉内には副結節を疑う複数の病変を認めます(cT4)。	左肺下葉中枢側にはsatellite lesionと思われる結節を複数認めます。
腫瘤は胸膜と広範に接しており、胸膜外脂肪層も一部途絶が見られ、壁側胸膜浸潤を疑います。	胸壁に広く接しており，胸壁脂肪織の消失も認められ，胸壁に浸潤している可能性もあります。
肋骨への浸潤ははっきりしません。 腫瘤は横隔膜とも一部接していますが、浸潤所見ははっきりしません。	
左肺門や気管分岐下や大動脈下などの同側の縦隔	左肺門部，気管分岐下，左下部気管傍リンパ節が

Subtasks for Radiology Report (RR)

(a) Named Entity Recognition (RR-NER)

Guideline Learning (proposed in Subtask 2 of NTCIR-16 Real-MedNLP)

- Provide only the guideline text for human annotators
  - 30-40 example sentences annotated
  - A sample of annotation guideline is available at https://sociocom.naist.jp/real-mednlp
- Participants can use any other resources outside this task if they find them useful for their methods
- Can we teach a model as if it is a human?



2.3.3 Features and Measurements

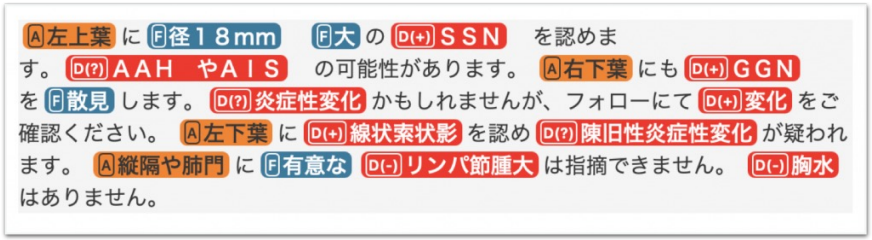
Features and Measurements tags are given to modifying phrases or predicative adjectives (such as "scattered") that pertain to the features, measurements, values, areas, or degrees of a given Disease or Symptom entity. They are also given to expressions that indicate degrees (such as "mild"). However, if the degree expression is connected with an expression that indicates changes, the degree expression is given a Change tag (<c>), (mentioned below in Section 2.3.4]), instead of features and measurements tags.

XML tag

<f>...</f>

Examples

(27) <f>Well-defined, smooth-margin</f> <d certainty="positive">nodular shadow</d> was recognized.



- Diseases and symptoms <d>
- Anatomical entities <a>
- Features and measurements <f>
- Change <c>
- Time <timex3>
- Test <t-test/key/val>
- Medicine <m-key/val>
- Remedy <r>
- Clinical Context <cc>

(b) TNM staging (RR-TNM)

To evaluate the generalization ability of NLP models when classifying radiology reports under multiple criteria (cancer staging)

According to predefined criteria, participant systems assign four labels

- tumor (T) 原発のがんの広がり
- lymph node (N) がん細胞のリンパ節への転移の有無と広がり
- metastasis (M) 原発から離れた臓器への遠隔転移
- clinical stage (cStage)

e.g.,

- Early cancers will be labeled with small numbers such as T1N0M0, cStage 1
- Advanced ones will be labeled with large numbers such as T3N2M1, cStage 4

FAQ for Radiology Report Subtasks

(a) How to build dataset

Answer: We collaborate with a teleradiology company and use an open-access radiology case repository. Radiology reports are written by radiologists diagnosing real, public, and anonymized cancer images.

(b) The dataset size

Answer: We are working on adding data, which will make the current dataset (n=135) larger.

(c) RR-TNM: definition of cancer stage

Answer: We follow the Union for International Cancer Control (UICC) criteria as clinicians do. We are planning to use the 7^{th} and 8^{th} editions to simulate revision.

Collaboration

To develop high-quality multilingual data, we collaborate with researchers from DFKI (Germany), LISN-CNRS (France), and RIKEN and NII (Japan)

- Akiko Aizawa, Ph.D. (NII, Japan)
- Gabriel Herman Bernardim Andrade (NAIST, Japan)
- Cyril Grouin, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)
- Shouhei Hanaoka, MD, PhD (The University of Tokyo, Japan)
- Thomas Lavergne, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)
- Yuji Matsumoto, Ph.D. (RIKEN, Japan)
- Faith Wavinya Mutinda (NAIST, Japan)
- Aurélie Névéol, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)
- Noriaki Nishida, Ph.D. (RIKEN, Japan)
- Tomohiro Nishiyama (NAIST, Japan)
- Patrick Paroubek, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)
- Lisa Raithel (DFKI, Germany)
- Roland Roller, Ph.D. (DFKI, Germany)
- Hiroki Teranishi, Ph.D. (RIKEN, Japan)
- Philippe Thomas, Ph.D. (DFKI, Germany)
- Narumi Tokunaga (RIKEN, Japan)
- Hui-Syuan Yeh (Université Paris-Saclay, CNRS, LISN, France)
- Pierre Zweigenbaum, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)

Tentative Schedule

- March 2023: Dataset Release
- (March-June 2023: Dry Run)
- June 1, 2023: Registration Deadline
- July 2023: Formal Run
- August 1, 2023: Evaluation Result Release
- August 1, 2023: Task overview paper release (draft)
- September 1, 2023: Submission due of participant papers (draft)
- November 1, 2023: Camera-ready participant paper due
- December 2023: NTCIR-17 Conference in NII, Tokyo