



MedNLPDoc

The 3rd Round of MedNLP Task

Aramaki E^{1,2}, Kano Y^{2,3}, Ohkuma T⁴, Morita M⁵

¹Kyoto Univ, ²JST PRESTO, ³Shizuoka Univ, ⁴Fuji Xerox, ⁵Univ of Tokyo

Background & Goal

- **Background**

- Medical records could potentially be rich information resource
- BUT: Medical NLP in Japan is not so powerful when compared to oversea countries



- **Goal of MedNLP Task**

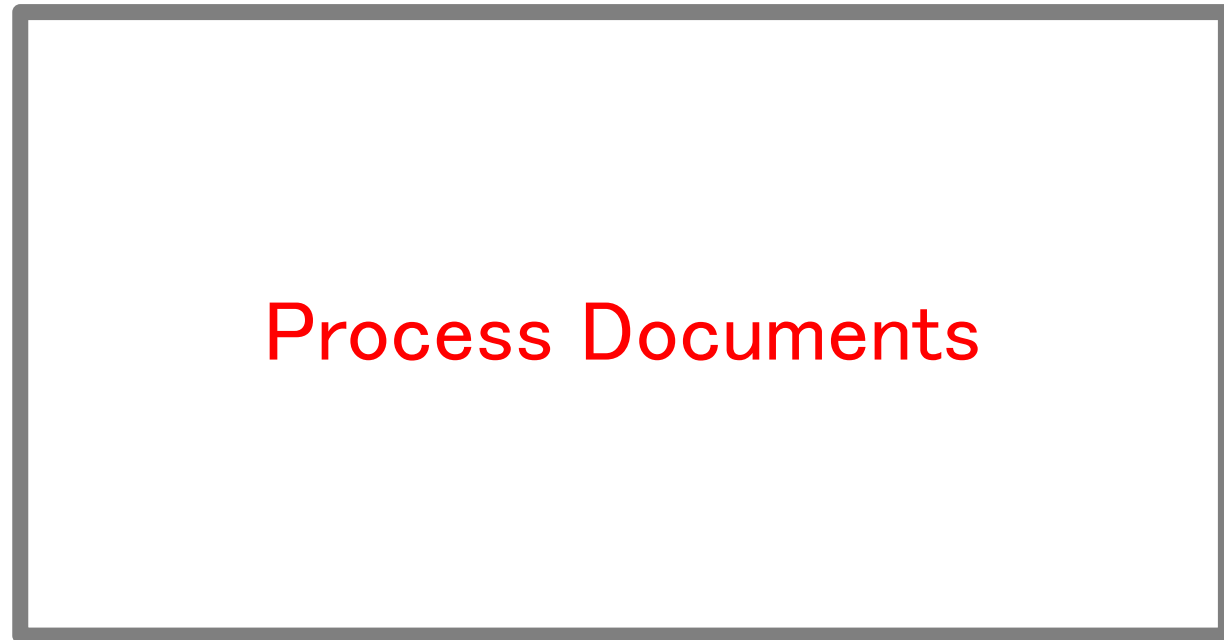
- To improve language processing techniques in medicine (in Japanese)
- To build community who are interested in this field

Strategy

IN

Medical Documents

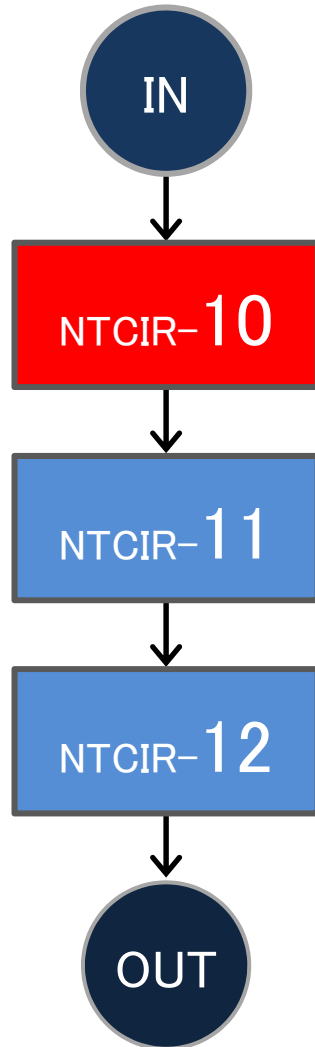
“ACTOS was prescribed at St Luke’s Hospital, and his leg was swollen.”



OUT

Better Clinical Solutions

Strategy



Medical Documents

“ACTOS was prescribed at St Luke’s Hospital, and his leg was swollen.”

De-identification

“ACTOS was prescribed at ■■■ Hospital, and his leg was swollen.”

NER (Named Entity Recognition)

“ACTOS was prescribed at ■■■ Hospital, and his leg was swollen.”

Normalization

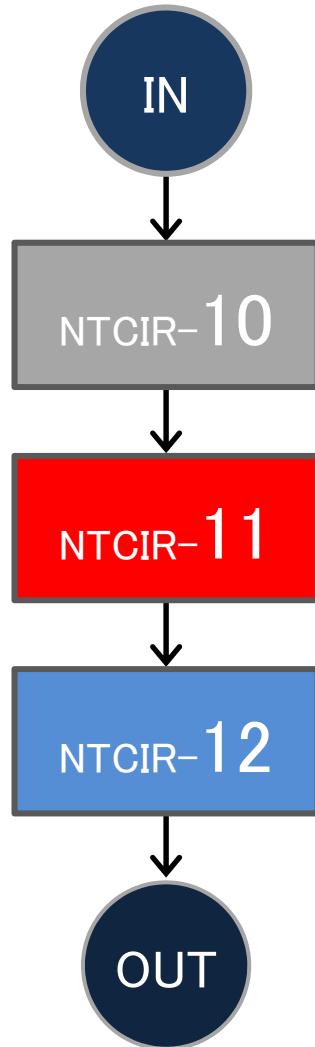
“Pioglitazone was prescribed at ■■■ Hospital, and his leg was edema.”

Decision Support

“Ψ(θ)_(θ)Ψ > ↑ \$*σ↗?”

Better Clinical Solutions

Strategy



Medical Documents

“ACTOS was prescribed at St Luke’s Hospital, and his leg was swollen.”

De-identification

“ACTOS was prescribed at ■■■ Hospital, and his leg was swollen.”

NER (Named Entity Recognition)

“ACTOS was prescribed at ■■■ Hospital, and his leg was swollen.”

Normalization

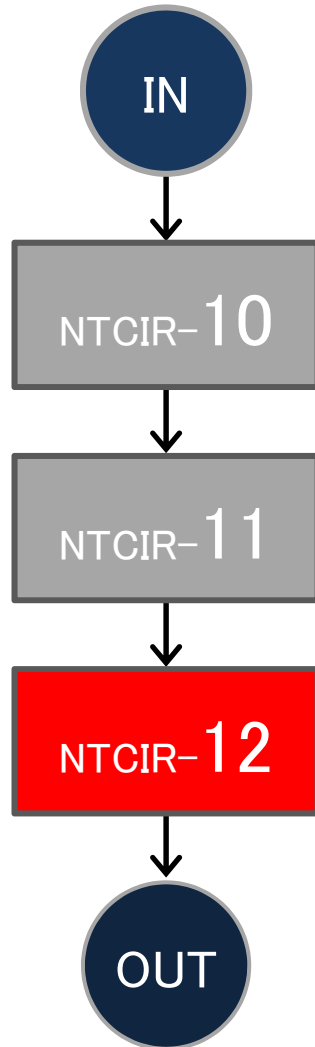
“Pioglitazone was prescribed at ■■■ Hospital, and his leg was edema.”

Decision Support

“Ψ(θ)_(θ)Ψ > ↑ \$※♂☞?”

Better Clinical Solutions

Strategy



Medical Documents

“ACTOS was prescribed at St Luke’s Hospital, and his leg was swollen.”

De-identification

“ACTOS was prescribed at ■■■ Hospital, and his leg was swollen.”

NER (Named Entity Recognition)

“ACTOS was prescribed at ■■■ Hospital, and his leg was swollen.”

Normalization

“Pioglitazone was prescribed at ■■■ Hospital, and his leg was edema.”

Decision Support

“Ψ(θ)_(θ)Ψ > ℓ \$※♂☞?”

Better Clinical Solutions

Example

Input Medical Document (in Japanese)

2022年10月3日に心窩部痛を自覚し、近藤医院を受診。
精査のために行った上部消化管内視鏡検査で胃体部
大弯に潰瘍性病変を認め、同部位を生検した。病理
組織診断によって腺癌を認め、胃癌の診断で精査加
療目的に12月1日当科入院。

NTCIR-10 MedNLP-1

Task 1: De-identification task

Task 2: Extracting complaint & diagnosis task

2022年10月3日に心窩部痛を自覚し、近藤医院を受診。

精査のために行った上部消化管内視鏡検査で胃体部

大弯に潰瘍性病変を認め、同部位を生検した。病理

組織診断によって腺癌を認め、胃癌の診断で精査加

療目的に12月1日当科入院。

NTCIR-11 MedNLP-2

Task 1: Extracting complaint & diagnosis task

Task 2: Normalizing complaint & diagnosis task

ICD-10: R101

2022年10月3日に**心窩部痛**を自覚し、近藤医院を受診。

精査のために行った上部消化管内視鏡検査で胃体部

大弯に潰瘍性病変を認め、同部位を生検した。病理

ICD-10: C16

ICD-10: C169

組織診断によって**腺癌**を認め、**胃癌**の診断で精査加

療目的に12月1日当科入院。

※ **ICD-10**: International Classification of Diseases, 10th Edition (WHO)

NTCIR-12 MedNLPDoc

Task 1: ICD-10 coding task

Coding to this **case**, NOT to **terms** in this document

2022年10月3日に心窩部痛を自覚し、近藤医院を受診。
精査のために行った上部消化管内視鏡検査で胃体部
大弯に潰瘍性病変を認め、同部位を生検した。病理
組織診断によって腺癌を認め、胃癌の診断で精査加

→**Diagnosis:** **ICD-10: C169**
Gastric cancer

Material: Text Book



Supported by

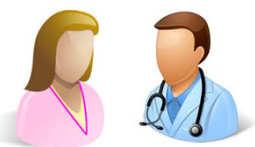


医学書院

編集：鳥羽 克子 / 診療情報管理東京ネットワーク

- 判型 B5
- 頁 356
- 発行 2006年04月
- 定価 3,888円 (本体3,600円+税8%)
- ISBN978-4-260-00219-6

Training book for
health information manager (HIM)
診療情報管理士



Doctor

Input=patient info. (face2face)



output=health record + disease name



HIM

Input=health record



output=disease name for insurance

Let's Try!

<data id="66" sex="m" age="45">

<text type="既往歴">なし

2005年1月 1月初旬から咳が続き売薬購入するも改善なし。

2月3日 他院受診。喀痰からG6号検出。

25日 肺結核の診断にて当院紹介入院。

INH 0.4, RFP 0.4, SM 1g/日、PZA 1.5g/日で化学療法スタート。

</text>

<icd code="A150">肺結核</icd>

Let's Try!

```
<data id="66" sex="m" age="45">  
<text type="既往歴">なし  
2005年1月 1月初旬から咳が続き売薬購入するも改善なし。  
2月3日 他院受診。喀痰からG6号検出。  
25日 肺結核の診断にて当院紹介入院。  
INH 0.4, RFP 0.4, SM 1g/日、PZA 1.5g/日  
で化学療法スタート。  
</text>
```

```
<icd code="A150">肺結核</icd>
```

Let's Try!

```
<data id="68" sex="m" age="49">
```

```
<text>
```

2004年12月2～16日、前回入院。

今回2回目の入院。

前回他院にてアメーバ肝膿瘍の手術予定であったが、術前の検査でHIV陽性であったため当院入院。

STS陽性。

今回上記の疾患について外来フォロー中であったが吐血で入院。

食道潰瘍が判明。

```
</text>
```

```
<icd code="K221">食道潰瘍</icd>
```

```
<icd code="D62">急性出血後貧血</icd>
```

```
<icd code="R75">HIV陽性</icd>
```

```
<icd code="A530">STS陽性</icd>
```

```
<icd code="Z861">アメーバ肝膿瘍</icd>
```

```
<icd code="K270">吐血</icd>
```

Let's Try!

```
<data id="68" sex="m" age="49">
```

```
<text>
```

2004年12月2～16日、前回入院。

今回2回目の入院。

前回他院にてアメーバ肝膿瘍の手術予定であったが、術前の検査でHIV陽性であったため当院入院。

STS陽性。

今回上記の疾患について外来フォロー中であったが吐血で入院。

食道潰瘍が判明。

```
</text>
```

```
<icd code="K221">食道潰瘍</icd>
```

```
<icd code="D62">急性出血後貧血</icd>
```

```
<icd code="R75">HIV陽性</icd>
```

```
<icd code="A530">STS陽性</icd>
```

```
<icd code="Z861">アメーバ肝膿瘍</icd>
```

```
<icd code="K270">吐血</icd>
```

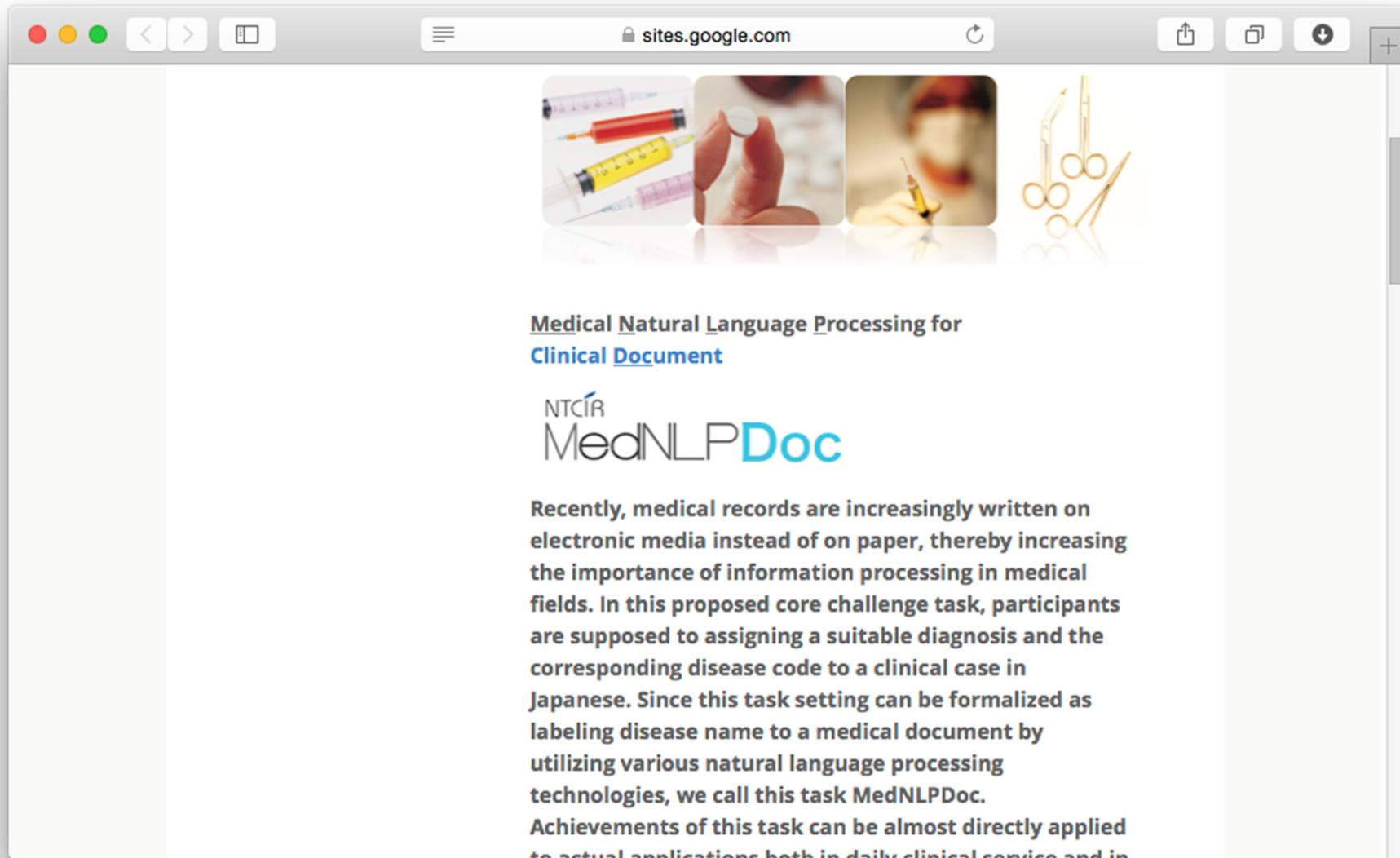
What is this task?

~ Possible Approaches ~

- **Multi Labeling for documents**
 - Large label space (more than 20,000 labels)
- **Document Similarity**
 - Find similar document & use their code (based on K-means...)
- **Rule based approach might still competitive**

I'm looking forward to see your idea

To get detailed information, Google "*mednlpdoc*"



The screenshot shows a web browser window with the address bar displaying "sites.google.com". The page content includes a header image with four panels: syringes, a hand holding a pill, a person with a needle, and surgical instruments. Below the image is the title "Medical Natural Language Processing for Clinical Document" and the logo "NTCIR MedNLPDoc". The main text describes the task of assigning a diagnosis and disease code to a clinical case in Japanese using NLP technologies.

Medical Natural Language Processing for Clinical Document

NTCIR
MedNLPDoc

Recently, medical records are increasingly written on electronic media instead of on paper, thereby increasing the importance of information processing in medical fields. In this proposed core challenge task, participants are supposed to assigning a suitable diagnosis and the corresponding disease code to a clinical case in Japanese. Since this task setting can be formalized as labeling disease name to a medical document by utilizing various natural language processing technologies, we call this task MedNLPDoc. Achievements of this task can be almost directly applied to actual applications both in daily clinical service and in



Better NLP, Better Medicine

MedNLP

- Medical Natural Language Processing



Handling medical records

Let's Try!

```
<data id="74" sex="f" age="74">
```

```
<text>
```

2004年3月20日 嘔気、腹痛にて発見された直腸がんに対して、前方切除術施行。

30日 術後経過良好にて退院。

12月24日 座位時一過性脳虚血発作あり。

その後、呂律障害、右上肢の感覚以上を認めたため頭部CT施行したところ頭頂葉に低呼吸域を認め、脳転移の疑いにて入院。

2005年1月4日放射線治療(全脳照射:40Gy/20回)開始。

また転移巣の検索において肝転移、肺転移が発見された。

2月8日 放射線治療終了するが全身倦怠感強く、これ以上の積極的な治療を望まなかったため、緩和ケア病棟入室となる。

4月4日 午前7時45分、永眠。

```
</text>
```

```
<icd code="C793">脳転移</icd>
```

```
<icd code="Z850">直腸がんの既往</icd>
```

```
<icd code="Z904">直腸の後天性欠損</icd>
```

```
<icd code="Z980">直腸吻合状態</icd>
```

```
<icd code="Z510">放射線治療施行中</icd>
```

```
<icd code="C787,M8010/6">肝転移</icd>
```

```
<icd code="C780,M8010/6">肺転移</icd>
```

```
<icd code="Z515">緩和ケア</icd>
```