

Annotation Guidelines

Quy tắc chung

- Một thực thể cần phải được định danh rõ, không bị nhập nhằng, giúp người đọc hình dung được thực thể đó là gì, ở đâu, tên gọi là gì, không bị nhầm lẫn sang các thực thể khác
 - "Đại học Ngoại Thương cơ sở Hà Nội", "Khoa Khám Bệnh, Bệnh viện Bạch Mai", "Số 127, đường Trần Hưng Đạo", "tầng 3, Tòa nhà Trúc Bạch"
- Chỉ gán nhãn các câu liên quan đến bệnh nhân COVID-19 (công bố ca bệnh mới, lịch trình di chuyển, triệu chứng, bệnh nền, công bố tử vong)
- Chỉ gán các câu liên quan đến dịch COVID-19 ở Việt Nam.
- Nếu 2 hay nhiều thực thể overlap thì gán nhãn sao cho thực thể được tạo thành là dài nhất, bao gồm các thực thể thành phần bị overlap.
 - Những trường hợp về từ các vùng mà không phải vùng đã công bố có bệnh nhân COVID-19, không phải địa điểm được Bộ Y tế công bố, sẽ thực hiện tự theo dõi sức khỏe tại nhà trong vòng 14 ngày, lấy mẫu xét nghiệm khi có dấu hiệu bệnh như: sốt, ho, đau họng, khó thở, **giảm hoặc mất vị giác, khứu giác**.
=> gán cả cụm "**giảm hoặc mất vị giác, khứu giác**" là 1 thực thể
- Trường hợp thực thể đi kèm tên viết tắt
 - Trung tâm kiểm soát dịch bệnh TP HCM (HCDC) => gán nhãn cả 2 thực thể là: "Trung tâm kiểm soát dịch bệnh TP HCM" và "HCDC"

Nhãn: PATIENT_ID, NAME, GENDER, AGE, JOB, LOCATION, ORGANIZATION, DATE, SYMPTOM_AND_DISEASE, TRANSPORTATION

Giao diện chung của công cụ gán nhãn

PATIENT_ID i NAME n AGE a GENDER g JOB j

ORGANIZATION o LOCATION l DATE d SYMPTOM_AND_DISEASE s

TRANSPORTATION t

Trong số ca bệnh được Bộ Y tế mới công bố, nam "bệnh nhân 974", 33 tuổi, làm nghề thợ xây, từng tham dự đám giỗ và đám cưới.

1 / 1004

Giao diện gồm 2 thành phần chính:

- Nhãn (bao gồm màu sắc đại diện và phím tắt)
- Câu cần gán nhãn

Cách sử dụng: dùng chuột bôi đen đoạn text cần gán nhãn => sử dụng phím tắt tương ứng với nhãn cần gán.

Chú ý:

- Không gán các khoảng trắng, dấu câu ở cuối các thực thể.

PATIENT_ID

- Tại Việt Nam, bệnh nhân COVID-19 được định danh bằng số thứ tự. X: số thứ tự (1, 2, 3, 100, 200, ...).
 - Bệnh nhân X, Bệnh nhân thứ X, bệnh nhân số X, BN X, => chỉ gán "X" với nhãn PATIENT_ID
 - BNX => gán nhãn cho "BNX" là PATIENT_ID

Đáng chú ý, trong số 5 bệnh nhân này có bệnh nhân **592 *** là cụ bà L.T.H. (100 tuổi, trú xã Duy Trung, huyện Duy Xuyên, Quảng Nam).

- **Chỉ gán nhãn các số thứ tự ở cấp độ quốc gia (Việt Nam).**
 - "T.A. là bệnh nhân COVID-19 đầu tiên của Hải Phòng trong giai đoạn này của dịch, còn anh C. là bệnh nhân thứ 2 của Hà Nam." => Không gán nhãn "2" là PATIENT_ID vì đây là bệnh nhân được đánh số ở cấp độ địa phương
- Trường hợp liệt kê một khoảng gồm nhiều bệnh nhân liên tiếp :
 - BN990-993: gán nhãn riêng biệt "BN990" và "993" là 2 thực thể nhãn PATIENT_ID

Theo đó, 4 bệnh nhân mới (**BN990 *** - **993 ***) ở **Đà Nẵng ***, **Quảng Nam *** và **Hải Dương ***, có thông tin cụ thể như sau:.

NAME

NAME

q

Tên riêng của người bao gồm:

- Tên bệnh nhân (để bảo vệ quyền riêng tư, tên bệnh nhân COVID-19 thường được viết tắt):
 - "N.H.N", "N.H.N.", "T.", "T" (**chú ý phân biệt dấu chấm trong tên bệnh nhân và các dấu chấm cuối câu**)
- Tên những người có liên quan hay tiếp xúc với bệnh nhân:
 - Ngày 21/8, bệnh nhân có tiếp xúc với cô **A.**, là nhân viên siêu thị Điện máy Xanh Đà Nẵng. => gán nhãn "**A.**"
- Các danh xưng "**ông**", "**bà**", "**anh**", "**chị**", "**giám đốc**", "**chủ tịch**", ... **KHÔNG** nằm trong tên riêng
 - ông **H.H.T.** => chỉ gán nhãn "**H.H.T.**" với nhãn NAME

T.A. * là bệnh nhân COVID-19 đầu tiên của Hải Phòng trong giai đoạn này của dịch, còn anh **C. *** là bệnh nhân thứ 2 của Hà Nam.

Các trường hợp không gán nhãn:

- Không gán nhãn những người không liên quan trực tiếp đến lịch trình di chuyển hay không có liên hệ, không tiếp xúc với bệnh nhân
 - "Bác sĩ **Trương Hữu Khanh**, chuyên gia về bệnh truyền nhiễm tại TP HCM, phân tích "bệnh nhân 34" lây nhiễm nhiều còn do tiếp xúc gần với nhiều người trong giai đoạn bắt đầu triệu chứng bệnh." => không gán nhãn "**Trương Hữu Khanh**"
- Không gán nhãn tên người đại diện, đưa ra phát ngôn hay công bố của 1 tổ chức
 - "Sáng 12-8, ông **Trương Huyền Trường** - giám đốc Bệnh viện Lao và bệnh phổi Quảng Trị (là nơi điều trị các bệnh nhân nhiễm COVID-19 của tỉnh) - cho biết đến thời điểm này, bệnh nhân 862 đã bỏ ăn bữa thứ tư và một mực đòi về nhà." => không gán nhãn "Trương Huyền Trường"

GENDER

GENDER

g

- **Chỉ gán nhãn giới tính của bệnh nhân và những người liên quan trực tiếp (tiếp xúc) với bệnh nhân. Trong một câu, giới tính cần được gán với một đối tượng được định danh (có tên, có mã bệnh nhân):**
 - "Bệnh nhân 125 là **nữ** chuyên gia quốc tịch Nam Phi, 22 tuổi, trú tại quận 7, TP.HCM." => gán nhãn "nữ" là thực thể nhãn GENDER.
 -
- **KHÔNG** gán nhãn các danh xưng: "ông", "bà", "cháu gái", "cháu dâu"

Theo đó, 3 bệnh nhân COVID-19 mới của Hải Dương bao gồm Đ.Q.M. (nam *, 17 tuổi) và N.T.C (nữ *, 72 tuổi, cùng trú tại thôn Nhân Lữ, xã Cẩm Chế, huyện Thanh Hà) và V.T.L. (nữ *, 59 tuổi).

AGE

AGE

a

- **Chỉ gán nhãn giá trị tuổi của bệnh nhân và những người liên quan (tiếp xúc) (KHÔNG gán nhãn từ "tuổi"). Trong một câu, tuổi cần được gán với một đối tượng được định danh (có tên, có mã bệnh nhân)**
 - ""Bệnh nhân 100", nam, 55 tuổi, địa chỉ ở quận 8, TP HCM." => gán nhãn "55" là thực thể kiểu AGE
 - Trường hợp liệt kê một khoảng tuổi của bệnh nhân:
 - "5 bệnh nhân còn lại được ghi nhận từ số 1025 đến 1029 gồm hai nam, ba nữ, tuổi từ 20 đến 70, đều ở Mỹ An, Ngũ Hành Sơn." => gán nhãn riêng biệt "20" và "70" là hai thực thể kiểu "AGE"

"Bệnh nhân 100", nam, 55 * tuổi, địa chỉ ở quận 8, TP HCM.

Trong đó, bệnh nhân 931 * đến 933 * gồm hai nam * và một nữ *, độ tuổi từ 14 * đến 66 *, đều ở thành phố Hội An *, Quảng Nam *.

Trường hợp suy thận * còn lại là bệnh nhân nam, 45 tuổi, ở huyện Hòa Vang *.

JOB

JOB

z

Chỉ gán nhãn nghề nghiệp của bệnh nhân và các cá nhân có liên quan trực tiếp (tiếp xúc, gặp mặt, ở gần). Ngoài ra, những từ chỉ nghề nghiệp cần phải được gán với 1 cá nhân nhất định trong câu (có tên, có mã bệnh nhân).

"Bệnh nhân 35" là **nhân viên bán hàng** tại Siêu thị Điện máy Xanh ở quận Hải Châu, Đà Nẵng." => gán nhãn "**nhân viên bán hàng**" là thực thể kiểu JOB do là nghề nghiệp của bệnh nhân 35

"Bệnh nhân 35" là **nhân viên bán hàng** tại Siêu thị Điện máy Xanh ở quận Hải Châu, Đà Nẵng.

CHÚ Ý:

- "du học sinh", "học sinh", "sinh viên" cũng được xem là JOB
 - "Bệnh nhân 129 là nam, 20 tuổi, địa chỉ ở Nghĩa Tân, Hà Nội, là **du học sinh** tại Anh, nhập cảnh về Nội Bài ngày 20-3 trên chuyến bay VN0054." => gán nhãn "**du học sinh**" là thực thể nhãn JOB
- Chỉ gán nhãn "**chuyên viên**", "**chuyên gia**", nếu có ngành nghề cụ thể đi kèm
 - "Họ đều là nam, tuổi 42-55, là **chuyên gia dầu khí**." => gán nhãn "**chuyên gia dầu khí**" là thực thể nhãn JOB
 - "Bộ Y tế trước đó thông tin bệnh nhân 125 là nữ **chuyên gia** quốc tịch Nam Phi, 22 tuổi, trú tại quận 7, TP.HCM." => **KHÔNG** gán nhãn "**chuyên gia**" là JOB
- Không gán nhãn các dạng viết tắt, viết gọn của nghề nghiệp: "**nam sinh**" (là viết gọn của "**nam sinh viên**", "**nữ sinh**" (là viết gọn của "**nữ sinh viên**"))

Bệnh nhân **425** là **nữ** **điều dưỡng** **24** tuổi, làm việc ở Khoa Hồi sức tích cực.

=> gán "điều dưỡng" là JOB vì là từ chỉ nghề nghiệp của bệnh nhân 425

LOCATION

LOCATION

C

Các thực thể LOCATION (địa điểm) có thể bao gồm:

- Tên châu lục: "**châu Âu**", "**châu Á**"
- Tên quốc gia: "**Việt Nam**", "**Anh**", "**Mỹ**"

- Tên đơn vị hành chính của quốc gia: "**tỉnh Hải Dương**", "**thành phố Hà Nội**", "**quận Hoàn Kiếm**" (**gán nhãn cả các từ chỉ đơn vị hành chính: tỉnh, thành phố, quận, huyện, đường**)
- Tên các công trình xây dựng, công trình kiến trúc mang tính công cộng: cầu, đường, cảng, cửa khẩu, sân bay, nhà ga, trường học, bệnh viện, trạm y tế, chung cư, ký túc xá, doanh trại quân đội.
- Tên các địa điểm mang tính thương mại: nhà hàng, quán ăn, khách sạn, chợ, siêu thị
- Các thực thể nằm trong một thực thể lớn khác (về vị trí) => xem các thực thể bao trùm là một phần của thực thể nhỏ hơn. Trường hợp có thể xác định rõ được thực thể mà không cần địa danh bao trùm thì gán địa danh bao trùm là một thực thể LOCATION riêng biệt
 - "**Khoa Phổi, Bệnh viện Bạch Mai**": một khoa nằm trong một bệnh viện
 - "**12 đường Trần Hưng Đạo**": số nhà nằm trên một con đường
 - "Ngày 26-8, bệnh nhân vào **Bệnh viện Bạch Mai (Hà Nội)**, trong các lý do dẫn đến sốc nhiễm khuẩn, suy đa tạng, tử vong có lý do sốt xuất huyết và ngộ độc paracetamol." => gán nhãn hai thực thể "**Bệnh viện Bạch Mai**" và "**Hà Nội**" là hai thực thể LOCATION riêng biệt
- **Địa chỉ:**
 - Từng cấp bậc đơn vị hành chính là một thực thể riêng biệt: "Bệnh nhân 420", nữ, 71 tuổi, ở **phường Hòa Cường Nam, quận Hải Châu, thành phố Đà Nẵng**. => gán nhãn 3 thực thể: "phường Hòa Cường Nam", "quận Hải Châu", "thành phố Đà Nẵng" với nhãn LOCATION

"Bệnh nhân 420", nữ, 71 tuổi, ở **phường Hòa Cường Nam ***, **quận Hải Châu ***,
thành phố Đà Nẵng *.

- **Số nhà phải bao gồm cả tên đường để tránh bị nhập nhằng:** "Bệnh nhân 66", nữ 21 tuổi, trú tại **Chung cư Park View, số 107 Nguyễn Đức Cảnh, Phú Mỹ Hưng, quận 7**. => gán nhãn 4 thực thể: "**Chung cư Park View**"; "**107 Nguyễn Đức Cảnh**"; "**Phú Mỹ Hưng**"; "**quận 7**" với nhãn LOCATION
- "Từ số X đến số Y": "Từ nhà số **125** đến nhà **139 phố Trúc Bạch**, nơi vừa phát hiện bệnh nhân N.H.N. nhiễm COVID-19 bị phong tỏa nghiêm ngặt sáng 7-3." => gán nhãn "**125**" và "**139 phố Trúc Bạch**" là 2 thực thể nhãn LOCATION

CHÚ Ý:

- Chỉ gán nhãn các địa danh liên quan đến lịch trình di chuyển của bệnh nhân.
- Không gán nhãn quốc tịch.
- Trường hợp trường đại học thuộc một cấp đại học cao hơn thì gán nhãn riêng biệt cho từng cấp

- **Đại học Ngoại Ngữ - Đại học Quốc gia Hà Nội** => gán nhãn 2 thực thể riêng biệt "Đại học Ngoại Ngữ", "Đại học Quốc gia Hà Nội" (tùy ngữ cảnh mà gán nhãn ORGANIZATION hoặc LOCATION)
- Không gán nhãn các địa danh không được định danh rõ ràng:
 - Sau khi xuất viện, bệnh nhân sẽ được đưa vào các bệnh viện ở **Đà Nẵng** để tiếp tục điều trị bệnh nền. => chỉ gán "**Đà Nẵng**" là thực thể kiểu LOCATION, **KHÔNG** gán nhãn "**bệnh viện ở Đà Nẵng**"
- Trường hợp khoa, phòng, ban, hội... thuộc một tổ chức, khu vực thì chỉ gán nhãn ORG khi có đầy đủ cả tên của tổ chức, khu vực
 - Nếu được ngăn cách bằng dấu phẩy, dấu gạch ngang: xem các thành phần tạo nên và dấu ngăn cách là 1 thực thể

Bệnh nhân là vợ đang chăm sóc ca bệnh **N.V.M. *** đang điều trị tại

khoa hồi sức tích cực - chống độc, Bệnh viện Đà Nẵng * từ ngày **6-7 *** đến nay.

- Nếu được ngăn cách bằng các từ "của", "thuộc", "tại": tách các thành phần tạo nên thành các thực thể riêng biệt

Chiều nay **10-4 ***, **nữ *** **điều dưỡng *** **B.T.H. *** (bệnh nhân số **86 ***), **54 ***

tuổi, điều dưỡng của **Phòng khám ngoại trú HIV - Trung tâm Bệnh nhiệt đới *** thuộc

Bệnh viện Bạch Mai *, cùng 8 bệnh nhân khác điều trị COVID-19 ở

Bệnh viện Nhiệt đới trung ương * (**cơ sở 2 ***, **Đông Anh ***, **Hà Nội ***) được

công bố khỏi bệnh.

ORGANIZATION

ORGANIZATION

r

Các thực thể kiểu ORGANIZATION (tổ chức) có thể bao gồm:

- Tên các cơ quan chính phủ: bộ ngành, uỷ ban nhân dân.

- Chỉ gán nhãn tên các cơ quan ở cấp độ địa phương (như UBND, Sở Y tế) **NẾU** có tên địa phương cụ thể đi cùng: "UBND Hà Nội", "Sở Y Tế Quảng Ninh"
- Gán nhãn tên viết gọn của các cơ quan ở cấp độ quốc gia: "**Bộ Y Tế**", "**Văn phòng Chính phủ**" do có thể ngầm hiểu "**Bộ Y Tế**" là viết tắt cho "**Bộ Y Tế Việt Nam**"
- Tên các cơ quan liên quan tới việc xử lý dịch tễ: "**Cơ quan kiểm soát dịch bệnh Hà Nội**", "**HCDC**", "**Viện Pasteur TP Hồ Chí Minh**".

Trưa 20-8, **Trung tâm Kiểm soát bệnh tật TP.HCM * (HCDC *)** phát thông báo

đơn vị này vừa phát hiện một trường hợp tái dương tính với virus SARS-CoV-2. Đó là bệnh nhân T.Đ.C. (nam, 27 tuổi, ngụ tại phường Hòa Thạnh, quận Tân Phú).

- Tên các công ty, tổ chức nơi bệnh nhân làm việc: "**Công ty Samsung**", "**Công ty Trường Sinh**".
 - Trường hợp nơi bệnh nhân làm việc là các địa điểm công cộng, thương mại (như đã được nêu trong mục **LOCATION**) thì gán nhãn là **LOCATION**.
 - Ngoại trừ các trường hợp thực thể đi kèm với những từ chỉ chức vụ của một tổ chức (giám đốc, hiệu trưởng, trưởng khoa) thì gán nhãn **ORGANIZATION**.
 - Bệnh nhân 198, 53 tuổi, nhân viên **công ty Trường Sinh**. => gán nhãn "**công ty Trường Sinh**" là ORGANIZATION.
 - Ca bệnh 669 (bệnh nhân 669): nam, 50 tuổi, có địa chỉ tại phường Quang Vinh, TP Biên Hòa, tỉnh Đồng Nai, liên quan Bệnh viện Đà Nẵng, là bác sĩ tại **Bệnh viện Đa khoa Đồng Nai**, thăm người thân tại khoa nội Bệnh viện Đà Nẵng, tiếp xúc gần bệnh nhân 510. => gán nhãn "**Bệnh viện Đa khoa Đồng Nai**" là LOCATION do là địa điểm có tính công cộng
 - Theo bác sĩ Ngô Đức Tuấn - giám đốc **Bệnh viện Đa khoa Đồng Nai**, bệnh nhân 669 là bác sĩ L.Đ.N. => gán nhãn "**Bệnh viện Đa khoa Đồng Nai**" là ORGANIZATION do đi kèm từ chỉ chức vụ của tổ chức là "giám đốc".
 - Anh là chồng của "bệnh nhân 981", trú ở TP Hội An, công tác tại **trường THPT Nông Sơn**, huyện Nông Sơn và coi thi tại trường THPT Khâm Đức, huyện Phước Sơn trước khi được xác định mắc Covid-19. => nhãn '**trường THPT Nông Sơn**' là LOCATION do là địa điểm có tính công cộng

Chú ý :

- Xem cả địa danh nơi tổ chức đặt tại là một phần của thực thể: **UBND Thành phố Hà Nội**.

- Chỉ gán nhãn những tổ chức có liên quan đến bệnh nhân: đơn vị công bố ca bệnh, đơn vị công bố khỏi bệnh, cơ quan làm việc của bệnh nhân, đơn vị có trách nhiệm xử lý ca bệnh (họp bàn đưa ra hướng giải quyết, phun thuốc, điều tra dịch tễ).
- Không gán nhãn tên các cơ quan báo chí, truyền thông (Tuổi Trẻ Online, VnExpress, Thanh Niên) chỉ có nhiệm vụ dẫn lời, tường thuật, công bố bệnh.
- Cần chú ý tới ngữ cảnh để xác định một thực thể có phải là LOCATION hay không (tránh nhập nhằng với ORGANIZATION):**
 - Thực thể kiểu ORGANIZATION phải là tổ chức bao gồm một hay nhiều cá nhân và phải có chức năng, công việc nhất định. Thực thể kiểu ORGANIZATION thường đóng vai trò là chủ ngữ, thực hiện một hành động nào đó trong câu.
 - Nếu thực thể được dùng chỉ để chỉ nơi chốn thì gán nhãn là LOCATION. Thực thể LOCATION thường có các giới từ như "tại", "ở", "đến" đi kèm (trừ các trường hợp là tên công ty, đơn vị làm việc của bệnh nhân thì gán nhãn là ORGANIZATION).
 - Trường hợp tên tổ chức được viết gọn "TP Hà Nội công bố lịch trình di chuyển của ca nhiễm thứ 17" (viết gọn cho "UBND TP Hà Nội công bố lịch trình di chuyển của ca nhiễm thứ 17") => gán nhãn "TP Hà Nội" là LOCATION

Thực thể	LOCATION	ORGANIZATION
Bệnh viện, trung tâm y tế	Sáng 13-8, bệnh nhân (BN) 716 và 719 được công bố khỏi bệnh tại Bệnh viện Bạch Mai .	Sáng 13-8, Bệnh viện Bạch Mai đã tổ chức công bố khỏi bệnh và xuất viện đối với bệnh nhân (BN) 716 và 719.
	Theo bác sĩ Ngô Đức Tuấn - hiện công tác tại Bệnh viện Đa khoa Đồng Nai , bệnh nhân 669 là bác sĩ L.Đ.N. - là bác sĩ của khoa ung bướu của bệnh viện. => " Bệnh viện Đa khoa Đồng Nai " là nơi làm việc nhưng là địa điểm mang tính công cộng => gán nhãn LOCATION	Theo bác sĩ Ngô Đức Tuấn - giám đốc Bệnh viện Đa khoa Đồng Nai , bệnh nhân 669 là bác sĩ L.Đ.N. - trưởng khoa ung bướu của bệnh viện. => "giám đốc", "trưởng khoa" là từ chỉ chức vụ của 1 tổ chức nên " Bệnh viện Đa khoa Đồng Nai " trong ngữ cảnh này được xem như một tổ chức => gán nhãn ORGANIZATION

DATE

DATE

d

- Ngày trong tiếng Việt thường có dạng X/Y, X-Y (X là ngày, Y là tháng) (Gán nhãn cả năm nếu có)
 - Sau khi trở về địa phương, bệnh nhân 61 từng làm chủ hôn hai đám cưới và dự thánh lễ tại xã Phước Nam trước khi nhập viện điều trị vào ngày **15-3**.

Sau khi trở về địa phương, bệnh nhân 61 từng làm chủ hôn hai đám cưới và dự thánh lễ tại xã Phước Nam trước khi nhập viện điều trị vào ngày **15-3**.

- Các trường hợp chỉ khoảng thời gian: gán nhãn riêng biệt ngày bắt đầu và kết thúc
 - Từ ngày **1-6/8**, bệnh nhân được chuyển lên cách ly và điều trị tại Trung tâm Y tế huyện Hòa Vang, với chẩn đoán khi vào viện: Viêm phổi do nCoV trên bệnh nhân đái tháo đường type 2, tăng huyết áp, biến chứng nhiễm trùng huyết từ đường niệu, suy đa tạng. => 2 entities kiểu DATE là : "1" và "6/8"

SYMPTOM AND DISEASE

SYMPTOM_AND_DISEASE

s

- Triệu chứng liên quan tới bệnh nhân COVID-19
 - "Khoảng 4-5 ngày sau tôi **sốt, ho, đau họng dữ dội, xét nghiệm dương tính**", bệnh nhân chia sẻ. => Có 3 entities loại SYMPTOM AND DISEASE là: "sốt", "ho", "đau họng dữ dội".
- Các loại bệnh khác mà bệnh nhân COVID-19 mắc phải:
 - "Bệnh nhân 737", nữ, 47 tuổi, địa chỉ quận Sơn Trà, thành phố Đà Nẵng, tiền sử **suy thận mạn giai đoạn cuối** đã chạy thận nhân tạo chu kỳ và đặt stent, **suy tim, tăng huyết áp**. => có 3 entities là "suy thận mạn giai đoạn cuối", "suy tim", "tăng huyết áp".

Ngày 12-8 * bệnh nhân tử vong lúc 19h, với chẩn đoán tử vong COVID-19 biến chứng viêm phổi nặng * trên bệnh nhân suy thận mãn giai đoạn cuối *, tăng huyết áp *, đái tháo đường type 2 *.

Chú ý:

- Gán nhãn cả những từ bỏ ngữ cho bệnh như:
 - Từ chỉ mức độ nghiêm trọng ("**mãn tính**", "**giai đoạn cuối**", "**không phục hồi**", "**không đáp ứng thuốc**", "**sốt 38 độ C**")
- Không gán nhãn các tác nhân gây bệnh ("**viêm phổi do COVID-19**", "**viêm phổi nặng do virus**")
- Không gán nhãn từ "**biến chứng**"
- Không gán nhãn các bệnh, triệu chứng có kèm từ phủ định phía trước ("**không ho**", "**không sốt**")
- Không gán nhãn phương pháp điều trị: "**chạy thận nhân tạo**", "**đặt stent**"
- Những trường hợp về từ các vùng mà không phải vùng đã công bố có bệnh nhân COVID-19, không phải địa điểm được Bộ Y tế công bố, sẽ thực hiện tự theo dõi sức khỏe tại nhà trong vòng 14 ngày, lấy mẫu xét nghiệm khi có dấu hiệu bệnh như: sốt, ho, đau họng, khó thở, **giảm hoặc mất vị giác, khứu giác**. => gán cả cụm "**giảm hoặc mất vị giác, khứu giác**" là 1 thực thể nhãn SYMPTOM_AND_DISEASE

TRANSPORTATION

TRANSPORTATION

t

- Phương tiện, cách thức di chuyển của bệnh nhân.
- Chỉ gán nhãn những loại phương tiện có xác định rõ số hiệu.
- Chỉ gán nhãn biển số, số hiệu của loại phương tiện di chuyển, không gán nhãn loại phương tiện di chuyển
 - "Bệnh nhân 110", nữ, 19 tuổi, ở Đống Đa, Hà Nội, là du học sinh tại Mỹ, quá cảnh tại Nhật Bản, về Hà Nội trên chuyến bay **JL751**, số ghế 1A ngày 19/3.
 - Sau khi xuống sân bay, 15h30 cùng ngày, T. đi xe khách Công Tạo mang biển số **51B-142.48** về huyện Bình Đại.

"Bệnh nhân 110", nữ, 19 tuổi, ở Đống Đa, Hà Nội, là du học sinh tại Mỹ, quá cảnh tại Nhật Bản, về Hà Nội trên chuyến bay JL751 ✖, số ghế 1A ngày 19/3.

Các trường hợp khác cần lưu ý:

- Trường hợp có sự overlap giữa 2 thực thể thì xử lý theo quy tắc chọn thực thể dài hơn

Từ 25-7 ✖ đến 27-7 ✖ bệnh nhân có đến quán bún bà P.T.A. ✖ (thôn Võng Nhi ✖, xã Cẩm Thanh ✖), tiếp xúc với bà A ✖, ngày 1-8 ✖ đến Trạm y tế xã khai báo y tế.

Gán nhãn **"quán bún bà P.T.A."** (LOCATION) thay vì **"P.T.A"** (NAME). Không gán nhãn **"Trạm y tế xã"** do cần suy ra từ ngữ cảnh của câu để xác định rõ là xã nào.

- Trường hợp có dạng "X ở các Y1, Y2, Y3..." với X, Y là các LOCATION hoặc ORGANIZATION (Y thường là địa danh hoặc tổ chức mà X đặt tại) thì chỉ gán nhãn các địa danh bao trùm (Y)

Từ ngày 20 ✖ đến 26-7 ✖, chị H. ✖ họ ✖, sốt nhẹ ✖, đau họng ✖ nhưng vẫn nghĩ là cảm cúm bình thường nên đi làm bình thường ở trạm y tế các xã Vĩnh Hà ✖, Vĩnh Nam ✖, Vĩnh Trung ✖, Vĩnh Tú ✖, Vĩnh Thái ✖ và thị trấn Hồ Xá ✖ (huyện Vĩnh Linh ✖).

-

Trước đó, vào sáng 19-9 ✖, Bệnh viện Phổi Đà Nẵng ✖ đã chuyển hai bệnh nhân COVID-19 đã được công bố khỏi bệnh trước đó nhiều ngày là bệnh nhân số 416 ✖ và BN 478 ✖ về Bệnh viện Đà Nẵng ✖ để tiếp tục điều trị bệnh nền.

"Bệnh viện Phổi Đà Nẵng" đóng vai trò là chủ ngữ của câu, thực hiện hành động là chuyển hai bệnh nhân => gán nhãn ORGANIZATION. **"Bệnh viện Đà Nẵng"** là nơi tiếp nhận bệnh nhân, được dùng như một địa điểm => gán nhãn LOCATION

-

Theo thông tin từ **Bộ Y tế ***, hai bệnh nhân kể trên là bác gái và lái xe của gia đình bệnh nhân thứ **17 *** **N.H.N. ***.

-

Bốn người "ngoài **VN0054**" gồm **N.T.T. ***, **nữ *** **24 *** tuổi từ **Anh *** về nước ngày **9-3 *** trên máy bay thuê riêng, người nhà và lái xe của bệnh nhân số **17 ***, nam bệnh nhân 27 tuổi đi từ **Hàn Quốc *** về trên chuyến bay **VJ981 *** ngày **4-3 ***.

"**VN0054**" là số hiệu chuyến bay nhưng đi kèm với từ phủ định "**ngoài**" nên trong câu này, "**VN0054**" không phải là phương tiện di chuyển của bệnh nhân => không gán nhãn "**VN0054**" là TRANSPORTATION

-

Tối **6-3 ***, sau khi điều tra dịch tễ bệnh nhân **N.H.N. ***, cơ quan chức năng đã ghi nhận ông **T. *** cũng có dấu hiệu nghi ngờ mắc bệnh nên đã chuyển mẫu tới **Viện Vệ sinh dịch tễ trung ương *** để xét nghiệm.

"Viện Vệ sinh dịch tễ trung ương" là nơi tiếp nhận mẫu bệnh phẩm nên cũng có thể coi như là 1 LOCATION. Tuy nhiên đây không phải là 1 địa điểm mà bệnh nhân từng tới nên vẫn gán nhãn "Viện Vệ sinh dịch tễ trung ương" là ORGANIZATION

-

Nữ * **điều dưỡng *** **Đ.T.T.H. ***, một trong hai nữ điều dưỡng nhiễm bệnh, là **nhân viên đón tiếp *** tại khu vực sàng lọc bệnh nhân nghi mắc COVID-19, tuy nhiên thông tin từ **Bộ Y tế *** cho hay **điều dưỡng *** **H. *** và đồng nghiệp không phải lây bệnh từ bệnh viện, mà nguồn lây từ bên ngoài.

Gán nhãn các cả từ "**điều dưỡng**" (JOB) và "**H.**" (NAME) xuất hiện ở phía sau (dù có thể "điều dưỡng H." và "điều dưỡng Đ.T.T.H." là chỉ cùng một đối tượng)

-

Lãnh đạo **huyện Chương Mỹ *** cho biết trên địa bàn huyện có một người tiếp xúc với bệnh nhân thứ **17 ***, đó là chị **D.T.T. ***, đã tiếp xúc với bệnh nhân ngày **2 *** và ngày **4-3 ***.

"**Lãnh đạo huyện Chương Mỹ**" chưa được định danh rõ ràng => chỉ gán "**huyện Chương Mỹ**" là LOCATION

-

Cách ly 28 y bác sĩ, 17 bệnh nhân **Bệnh viện Thận Hà Nội *** Ông Nguyễn Khắc Hiền, giám đốc **Sở Y tế Hà Nội ***, cho biết từ tối **9-4 ***, 28 nhân viên y tế và 17 bệnh nhân của **Bệnh viện Thận Hà Nội *** đã được cách ly do tiếp xúc gần với bệnh nhân COVID-19 số **254 ***, đồng thời là bệnh nhân đang điều trị **suy thận mãn *** tại bệnh viện.

Không gán nhãn "y bác sĩ", "nhân viên y tế" do không có thể gắn với cá nhân cụ thể nào trong câu

-

Chiều nay **10-4 ***, **nữ *** **điều dưỡng *** **B.T.H. *** (bệnh nhân số **86 ***), **54 *** tuổi, điều dưỡng của **Phòng khám ngoại trú HIV - Trung tâm Bệnh nhiệt đới *** thuộc **Bệnh viện Bạch Mai ***, cùng 8 bệnh nhân khác điều trị COVID-19 ở **Bệnh viện Nhiệt đới trung ương *** (**cơ sở 2 ***, **Đông Anh ***, **Hà Nội ***) được công bố khỏi bệnh.

"Bệnh viện Nhiệt đới trung ương (cơ sở 2, Đông Anh, Hà Nội)": tách các thành phần bổ trợ trong ngoặc đơn thành các thực thể LOCATION

-

Trước đó, cũng trong ngày 7-3, Sở Y tế Cao Bằng đã có công văn khẩn số 741/SYT-NV đề nghị UBND hai huyện Quảng Uyên và Trùng Khánh yêu cầu người dân không di chuyển ra khỏi làng và nhân viên hai cơ sở du lịch nói trên không di chuyển ra khỏi nơi làm việc vì có thông tin 5 người nước ngoài đi cùng chuyến bay với bệnh nhân N.H.N. (người thứ 17 nhiễm COVID-19) đến các địa chỉ nói trên tại Cao Bằng để du lịch.

"UBND hai huyện Quảng Uyên và Trùng Khánh" ý muốn chỉ UBND huyện Quảng Uyên và UBND huyện Trùng Khánh. Trong trường hợp này, hai thực thể trên được dùng với cách nói gộp => chỉ gán nhãn các địa điểm là "huyện Quảng Uyên" (LOCATION) và "Trùng Khánh" (LOCATION)

-

"Bệnh nhân 435", nữ, 29 tuổi, ở đường Hoàng Hoa Thám, phường Vĩnh Trung, quận Thanh Khê, thành phố Đà Nẵng, nhân viên Khoa Y học cổ truyền, Bệnh viện Đà Nẵng.

Không gán "nhân viên" do không chỉ một nghề nghiệp cụ thể. "nhân viên" cũng không chỉ chức vụ trong tổ chức nên gán "Khoa Y học cổ truyền, Bệnh viện Đà Nẵng" là LOCATION

-

Tối 2/8, Thứ trưởng Bộ Y tế Nguyễn Trường Sơn, Trưởng Bộ phận thường trực đặc biệt chống dịch Covid-19 của Bộ Y tế tại TP Đà Nẵng, cho biết "bệnh nhân 429", nữ, 53 tuổi, tiền sử suy tim, suy thận mạn 5 năm rưỡi, đái tháo đường tuýp 2.

Trường hợp thực thể gồm nhiều thành phần được ngăn cách bởi các từ "của", "thuộc" => gán nhãn riêng biệt từng thành phần

-

Ngày 12-7 *, ông ra Bệnh viện Đà Nẵng * thăm mẹ là L.T.Đ. * (BN524 *) nằm ở khoa thận - tiết niệu tầng 6 (P606), sau đó về nhà, đến 20h đi chợ Nam Phước *, sau đó về nhà (có tiếp xúc một số người).

Không gán "khoa thận -tiết niệu tầng 6 (P606)" do cần suy ra từ ngữ cảnh của câu để biết khoa này thuộc "Bệnh viện Đà Nẵng"

•

Ngày 7/8 *, người này từ Đài Loan * về sân bay Cam Ranh * trên chuyến bay VJ2849 *, cách ly ngay tại Khánh Hòa *, mẫu bệnh phẩm gửi Viện Pasteur Nha Trang * xét nghiệm lần hai dương tính với nCoV.

"Viện Pasteur Nha Trang" là nơi tiếp nhận mẫu bệnh phẩm nên có thể coi như địa điểm. Đồng thời, "Viện Pasteur Nha Trang" cũng thực hiện chức năng là xét nghiệm mẫu bệnh phẩm nên có thể coi như tổ chức. Tuy nhiên, đây không phải là địa điểm mà bệnh nhân từng đến => gán "Viện Pasteur Nha Trang" với nhãn ORGANIZATION.